Psychological Review

[illegible]

Series of Monograph Supplements

Vol. III., No. [illegible] (Whole No. 16), [illegible]

The Correlation of Mental and Physical Tests

BY

CLARK WISSLER, A.M.

University Fellow in Psychology, Columbia University

[Submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in the Faculty of Philosophy, Columbia University. [illegible] No. 2 of Columbia University Contributions to Philosophy, Psychology [illegible].]

THE MACMILLAN COMPANY

[illegible] LANCASTER, PA.

[illegible] NEW YORK

[illegible]

THE

Psychological Review

EDITED BY

J. MARK BALDWIN
PRINCETON UNIVERSITY

J. McKEEN CATTELL
COLUMBIA UNIVERSITY

AND

H. C. WARREN
PRINCETON UNIVERSITY
(*Associate Editor and Business Manager*)

WITH THE CO-OPERATION OF

ALFRED BINET, ÉCOLE DES HAUTES-ÉTUDES, PARIS; JOHN DEWEY, H. H. DONALDSON, UNIVERSITY OF CHICAGO; G. S. FULLERTON, UNIVERSITY OF PENNSYLVANIA; G. H. HOWISON, UNIVERSITY OF CALIFORNIA; JOSEPH JASTROW, UNIVERSITY OF WISCONSIN; G. T. LADD, YALE UNIVERSITY; HUGO MÜNSTERBERG, HARVARD UNIVERSITY; M. ALLEN STARR, COLLEGE OF PHYSICIANS AND SURGEONS, NEW YORK; CARL STUMPF, UNIVERSITY, BERLIN; JAMES SULLY, UNIVERSITY COLLEGE, LONDON.

Series of Monograph Supplements,
Vol. III., No. 6 (Whole No. 16), June, 1901.

The Correlation of Mental and Physical Tests

BY

CLARK WISSLER, A.M.,
University Fellow in Psychology, Columbia University.

[Submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in the Faculty of Philosophy, Columbia University, and being Vol. IX., No. 2, of Columbia University Contributions to Philosophy, Psychology and Education.]

THE MACMILLAN COMPANY,
41 N. QUEEN ST., LANCASTER, PA.
66 FIFTH AVENUE, NEW YORK.
AGENT: G. E. STECHERT, LONDON (2 Star Yard, Carey St., W. C.); LEIPZIG (Hospital St., 10); PARIS (76 rue de Rennes).

CONTENTS.

PRESS OF
THE NEW ERA PRINTING COMPANY,
LANCASTER, PA.

INTRODUCTION.

This research is occasioned by the fact that psychologists and students of education have proposed certain tests, put them to trial and recorded their results, hoping thereby to find a means by which the fundamental elements of general and specific ability could be isolated and valued. In this they have not been without precedent, since upon the same assumption fitness for the civil and military service, for academic degrees and honors, for professional and technical licenses, etc., is usually determined by arbitrary tests and estimated by numerical averages in grade books. Our work is primarily with the grade book of the psychologist. It is of both theoretical and practical importance to know what relations exist between the results of his tests and those of others. The contention that all tests are arbitrary and futile has no weight in this connection so long as people go on using them. Educational, professional, physical and psychological tests are with us and bid fair to remain, at least until something better is found. Thus the times demand that the results obtained by the various tests be made an object of study. The most obvious line of approach in this problem is through correlation. If a test is general, then its results should correlate with many other special tests, and, in turn, if there is an integral relation between general and special ability, the results of the latter should correlate with the former. Should two tests show no correlation whatever, we can do no more than regard them as defining two entirely independent forms of activity. To determine the relative value of tests with respect to their general or specific significance, we must find some way of estimating the degree of correlation in terms of variability.

To the reader is due some explanation as to the methods of treating data. It is not proper to demonstrate these methods here, because they are not the objects of investigation, but, while they are treated at length in the appropriate literature, the mode of presentation is beyond the comprehension of all save the ex-

pert or those few who can afford to spend their time working up to them. So it seems necessary to give a few words to the methods employed.

The first thing to learn is that when mathematical formulas are resorted to it is only for purposes of convenient and exact statement. Long ago astronomers and engineers discovered that errors of observation are distributed in a certain orderly manner and are consequently susceptible to mathematical statement. By proper treatment they are able to estimate the precision of measurements, determine the number of observations necessary to a given standard of precision, etc. It was soon discovered that biological variation followed similar laws and that such variation could be measured accurately, thus enabling us to determine differences of race, type, species, etc. When psychologists took up their side of the question as to the precision of observations they soon discovered that all human performances, when objectified in units of space, time, etc., seem to follow certain laws of variability, and that these laws are in turn similar to those already worked out. In other words, variability is no longer a barrier to the study of human activity, because we can measure that variability. In more recent times it has been found possible to deal with variability in such a way that the functional and structural relations between phenomena can be accurately estimated, or, in technical terms, a method of correlation has been developed. As may be inferred, these methods are based upon empirical study, they have been attested by use and are the work of some of the best mathematicians and scientists of the past and present centuries. The reader who doubts their validity must look to the authors themselves.

In the following pages only such statements of the method have been offered as will enable the reader to understand the force of the results. In correlation it need only be borne in mind that we are using an accepted method to estimate the relative necessary relations between the phenomena under consideration. To make any such comparison at all we must assign our results a place in a scale of values. This is what the method of correlation does.

It remains for the writer to define his relation to this research. The tests were devised and conducted by Professor J. McKeen Cattell and his associates, but for the methods of compilation, together with all conclusions and opinions respecting the results and the validity of the various tests, the writer is responsible. The conception of the problem and the accumulation of material must be credited to the former, while the latter has only undertaken the compilation of results.

It is scarcely possible to mention the names of all who assisted in making the tests, but special note should be made of the fact that most of the tests on Barnard students were made by Miss S. L. Cody.

THE TESTS EMPLOYED.

These tests are made yearly upon sixty to seventy freshmen of Columbia College and repeated with those who remain to the end of the senior year. A preliminary report of methods and class averages, published in 1896,[1] gives in detail the methods of recording and preserving data. However, some recent additions and changes make it desirable to repeat them.

The tests now made in the psychological laboratory are as follows: length and breadth of head, strength of hands, fatigue, eyesight, color vision, hearing, perception of pitch, perception of weight, sensation areas, sensitiveness to pain, perception of size, color preference, reaction time, rate of perception, naming colors, rate of movement, accuracy of movement, perception of time, association, imagery, memory (auditory, visual, logical and retrospective). Records of stature, weight, etc., together with data concerning parentage, personal habits and health, are a part of the gymnasium tests required of all students in Columbia College.

At the beginning of the tests in the psychological laboratory a few words are said to the student concerning the object of the tests and the value of the results. He is then told to write his name, class and date of birth on the record blank; also to give information as to any tests already made upon him and necessary information as to how a copy of such measurements may be obtained. While the student is thus engaged the observer fills in the following:

What is his apparent age? (......), 17 (......), 18 (......), 19 (......), 20 (......), (......).

Is his apparent state of health good (......), above medium (......), below medium (......), poor (......)?

Is he tall (......), above medium (......), below (......), short (......)?

Is his head large (......), above medium (......), below medium (......), small (......)?

[1] Physical and Mental Measurements of the Students of Columbia University. Cattell and Farrand. PSYCHOLOGICAL REVIEW, Vol. III., p. 618.

Do you think his physical development good (......), above medium (......), below medium (......), poor (.....)?
Do you think him likely to be as a student good (......), above medium (......), below medium (......), poor (......)?
In these mental tests do you think him likely to be good (......), above medium (......), below medium (......), poor (......)?

Of this blank the student is kept in ignorance. He is asked whether he is right-handed for all purposes, the observer noting independently the preferred hand in writing and the other movements necessary in the tests.

Perception of Size.—This test is made first, in order that retrospective memory may be tested at the end of the series. A sheet of paper bearing a 5 cm. line is placed before the student and a blank sheet of the same size on the right. Without moving them or altering much his point of view he is required to draw a line as nearly as possible the length of the standard. This done he is requested to bisect the line drawn and from the middle draw a perpendicular as long as the whole line and then bisect the right-hand angle. The student is required to do this quickly.

Size of Head.—In measuring the head the maximum length and breadth are taken with calipers. For length the directions to observers are, to place one point of the calipers on the most prominent point of the forehead, between the eyebrows. If in this region very prominent bony ridges are felt, take a second measurement from a point just above the ridges and note both. Bring the other point of the calipers down to the posterior part of the head and move it along the middle line until the greatest length of head is found. For breadth of head take the maximum above the ears wherever found. Hold the calipers horizontally and perfectly symmetrically. Make two measurements of each diameter independently and record the results in mm. in the order in which they were made. Repeat until the average variation of the accordant measurements is about 1 mm. Leave a record of all measurements made.

Strength of Hand.—This test is made with the oval dynamometer. The student is shown how to hold the instrument and makes the test standing. He is not to see the dial or the record while making the test. The order is right-hand, left, right, left—four trials in all.

Fatigue.—The test is made with Cattell's ergometer. The pressure is applied by the ends of the index finger and thumb and the tension of the spring registered on a counter dial. The right hand is used and one trial of 50 efforts in a rhythm of about one a second is taken as the result. The observer records the reading for each ten efforts.

Eyesight.—These tests are made with Galton's instrument which gives the distance in cm. at which diamond numerals can be read by each eye singly. The right eye is tested first, beginning at a distance of 44 cm. From this the student proceeds until a card is reached where he can read at least eight out of ten numerals. The test is made in ordinary daylight.

Color Vision.—The subject is required to select the green shades from the woolen skeins supplied by the Cambridge Scientific Instrument Company in accordance with Mr. Galton's instructions.

Hearing.—The test is made in a quiet room with the ticking of a stop-watch as the standard. 5 m. and above is taken as the normal distance and under 1m. as abnormal.

Perception of Pitch.—The test is made with a monochord tuned so that F below middle C is given when the bridge is at 75 cm. The directions for making the test are as follows: "Give tone F twice at an interval of about two seconds to the student, whose back is turned. Then shift the bridge to about 50 and let the student find the tone. He must be warned against humming the tune and must probably be taught in advance how to use the monochord. Record the position of the bridge and then give the original tone twice and shift the bridge to the place where it was left by the student in his first trial, telling him that it is put back to this place. Let him now find the tone and record the position. Ask the student whether he plays a musical instrument or sings and record his answers."

Perception of Weight or Force of Movement.—In this test the lift is vertical and the dynamometer gives a pressure of 1 kg. to 10 cm. A mechanical stop is provided at a pressure of 1 kg. to give the student his standard. In making the test he is told to lift the handle to the stop three times and then make ten attempts to lift it to the same height after the operator has re-

moved the stop. Each lift is to be made in about 2 sec., with equal pauses between. A graphic record of the lifts is taken on a kymograph and filed with the other data.

Sensation Areas.—The points of the æsthesiometer are 2 cm. apart and the instrument is applied longitudinally to the back of the left hand, between the bones of the second and third fingers. Five tests are made, the student being touched with one or two points in the order, two, two, one, one, two, and being required to decide in each case whether he was touched with one or with two points.

Sensitiveness to Pain.—This is determined for the ball of the thumb of the right and left hands. An algometer is used in which the surface applied is of rubber, 1 cm. in diameter and rounded at the corners. The instrument is applied with gradually increasing pressure by the observer and the student is told to say as soon as the pressure becomes disagreeable. If he show signs of discomfort the pressure is stopped. Two tests are made on each hand in alternation, beginning with the right hand.

Color Preference.—Rectangles (5 x 3 cm.) of red, orange, yellow, green, blue, violet and white are shown in an irregular group on a black ground and the student asked to specify his likes and dislikes.

Reaction-time.—The reaction-time for sound is taken five times in succession with the Hipp chronoscope. After the reactions the student is asked whether he attended to the sound, to the hand or to both.

Rate of Perception.—A blank is provided containing 500 11-point capital letters, of which 100 are A's. Each of the other letters occurs 16 times and the whole series is arranged in an order drawn by lot. The student is required to mark as quickly as possible all the A's, the observer taking the time with a stop-watch. The blank is kept as a part of the student's test record and shows the accuracy of the performance as well as its quickness.

Naming Colors.—One hundred 1 cm. squares of colored paper (red, orange, yellow, green, blue, violet, pink, gray and black) arranged in chance order on a white ground are to be

named as quickly as possible. The observer takes the time with a watch and notes the errors. This is really a test in rate of reading or naming familiar things. Care is taken to see that all students have a ready name for each color on the card before taking the test.

Rate of Movement.—A blank ruled into one hundred 1-cm. squares into each of which the student must put a dot, completing the task as quickly as possible, constitutes this test. The observer records the time with a watch and preserves the record.

Accuracy of Movement. — Here a blank with 100 dots arranged in the form of a 10-cm. square is provided, the student being required to strike at each dot in succession, the aim being to hit them as nearly as he can and as quickly as possible. The time is taken by the observer and the blank preserved for the computation of accuracy.

Rhythm and Perception of Time.—The present test is of the ability to follow a given rhythm. The student makes with a telegraph key taps in the rhythm of sounds which occur one per second. He is told to continue tapping fifty times at the same rate after the sounder is stopped, which is after ten beats. The standard is given by a telegraph sounder operated by a clock. The student's tapping is recorded on a kymograph with a clock line in parallel.

Association.—A blank is provided containing the following words in bold-faced type: house, tree, child, time, art, London, Napoleon, red, enough. The observer explains the test to the student, and, when everything is ready, the blank is handed to the latter, who writes after each word as quickly as possible what it suggests to him, preferably a single word. The observer takes the time and files the blank.

Imagery.—The student answers the following questions:

Think of your breakfast table as you sat down to it this morning; call up the appearance of the table, the dishes and food on it, the persons present, etc. Then write answers to the following questions:

1. Are the outlines of the objects distinct and sharp?

2. Are the colors bright and natural?

3. Where does the image seem to be situated? In the head? Before the eyes? At a distance?

4. How does the size of the image compare with the actual size of the scene?

---

1. Can you call to mind better the face or the voice of a friend?

2. When "violin" is suggested, do you first think of the appearance of the instrument or the sounds made when it is played?

3. (*a*) Can you call to mind natural scenery so that it gives you pleasure? (*b*) Music? (*c*) The taste of fruit?

4. Have you ever mistaken an hallucination for a perception, *e. g.*, apparently heard a voice or seen a figure when none was present? If you answer "yes," describe the appearance on the back of this sheet.

Memory.—Four different tests are made in the order given here. The observer instructs the student as to what is expected of him in each case.

1. Auditory.—Each series of numerals in the following is read at a rate of about 2 per second, after which the student writes it from memory:

4	8	3	7	1	9	6	2
7	5	9	2	8	6	4	1
3	7	5	2	9	6	4	8

2. Visual.—Corresponding numerals are shown at the same rate.

3	9	5	2	6	8	1	4
8	5	2	7	4	6	3	1
5	1	6	2	7	3	4	8

3. Logical.—The following passage containing 100 words is read to the student, who then writes as much of it as he can. He is directed to give the words wherever possible, but attempt to give the thought completely.

"Tests such as we are now making are of value both for the advancement of science and for the information of the student who is tested. It is of importance for science to learn how people differ and on what factors these differences depend. If we can disentangle the complex influences of heredity and environment we may be able to apply our knowledge to guide human development. Then it is well for each of us to know in what way he differs from others. We may thus in some cases correct defects and develop aptitudes which we might otherwise neglect."

4. Retrospective.—At the beginning of the test a 5-cm. line was shown the student in the size test. At the end of the hour he is reminded of the line and requested to draw it from memory. This was devised as a test of memory for a thing which one had no special object in remembering.

During the test as opportunity may offer the observer fills in the following:

Forehead: straight (............), rather straight (............), somewhat sloping (............), sloping (............)?
Hair: black (............), dark brown (............), light brown (............), flaxen (............), red (............)?
Complexion: dark (...... .....), rather dark (............), rather light (.........), light (............)? clear(............), fairly clear (............), not clear (............), blotched (............)?
Eyes: gray (............), blue (............), brown (............)?
Hair: straight (............), slightly wavy (............), somewhat curly (............), curly (............)?
Nose: convex (............), slightly convex (............), slightly concave (............), concave (............)? Elevation: high (............), above medium (............), below medium (............), low (............)?
Ears: large (............), above medium (............), below medium (.........), small (............); projecting (............), somewhat projecting (.........), rather close (............), close (............)?
Mouth: large (............), above medium (............), below medium (............), small (............)?
Lips: thick (............), above medium (............), below medium (.........), thin (............)?
Hands (in relation to size of body): large (............), above medium (............), below medium (............), small (............)?
Fingers (in relation to width of hand): long (............), above medium (............), below medium (............), short (............)?
Face and Head: note asymmetry, also any abnormality, as malformation of ears, squint, etc.

After having discussed the subject the observer fills in the following:

Do you think his state of health good (............), above medium (............), below medium (............), poor (............)?
Do you think his physical development good (............), above medium (............), below medium (............), poor (............)?
Do you think him likely to be as a student good (............), above medium (............), below medium (............), poor (............)?
Do you think that in the mental tests he has done well (............), above medium (............), below medium (............), poorly (............)?
In understanding what was wanted, was he quick (............), above medium (............), below medium (............), slow (............)?
Was he talkative (............), above medium (............), below medium (............), quiet (............)?
Do you judge him to be accurate (............), above medium (............), below medium (............), not accurate (............)?
Do you judge him to be straightforward (............), above medium (.........) below medium (............), not straightforward (............)?

Do you judge him to be intellectual (............), above medium (............), below medium (............), not intellectual (............)?
Do you judge his will to be strong (............), above medium (............), below medium (............), weak (............)?
Do you judge his emotions to be strong (............), above medium (.........), below medium (............), weak (............)?
Would you call him well-balanced (............), above medium (............), below medium (............), not well-balanced (............)?
Would you call his temperament choleric (............), sanguine (............), melancholic (............), phlegmatic (............)?

The students are taken through the tests singly. Students and officers of the department act as observers and the average time required to make the test is 45 min. The above enumeration of tests is the order in which they occur on the blank, but as three freshmen are usually taken through during the same period and it being impossible to duplicate all the apparatus the regular order of the blank is not followed, and the tests are arranged in definite groups in separate rooms and each group given in the order of its occurrence in the blank. In all cases the size test is made first and the fatigue and retrospective memory tests at the end of the series. By keeping a record of the order of the three observers for each test period it is possible to divide the students into three groups in each of which all followed the same order.

It is proposed to add tests of attention, apperception, suggestion, etc., to the above series, but so far no very satisfactory method has been evolved. In case of apperception, supplying the missing words in the following was tried:

The manner in which..of imagery and ....follow each other through our.................................. the...flight of one idea before the next, the ....our minds make between things wide as the ....asunder, transitions which at first sight ....us by their abruptness, but which, when ....closely, often reveal...................................links of.......................................naturalness and...all this.., ...streaming has from time immemorial excited the...of all those ....happened to be caught by its................................... mystery.

This, like most mental tests, is in the nature of an experiment. No one can tell offhand what will make a good test for a particular function.

In the observer's blank it will be seen that certain ready estimates of the student are required. The data thus obtained is more for estimating the precision of such methods of observation than for affording information concerning the subject himself. In many cases it is possible to compare the judgment of the observer with actual measurements. By some such procedure it seems possible to determine the general probable error of ordinary judgments.

In connecton with the gymnasium measurements the student is required to supply the following information:

	Birthplace	Date of birth
Nationality	of father	of mother
	his father	her father
	his mother	her mother

	Father.	Mother.	Paternal Grandfather.	Paternal Grandmother.	Maternal Grandfather.	Maternal Grandmother.
Living? (if so, give age)						
Deceased? (if so, give year of death and age at time of death)						
Cause of death, if deceased						
Most serious diseases from which they have suffered						

	1	2	3	4	5	6	etc.
Your mother's children . . . . { born. deceased.							

[Write B for brother and S for sister in the order of age and in the proper column. Include yourself designated by X. After B, S or X write date of birth thus: B. Feb. 10, '84. In case any brothers or sisters have died, write date and cause of death after 'deceased.']

How many brothers did your father have? ( ), how many sisters? ( ) was your father his mother's 1st ( ), 2d ( ), 3d ( ), 4th ( ), 5th ( ), 6th ( ), or what ( ) child?

How many brothers did your mother have? ( ), how many sisters? ( ), was your mother her mother's 1st ( ), 2d ( ), 3d ( ), 4th ( ), 5th ( ), 6th ( ), or what ( ) child?

Which side of the family do you most resemble in physical build?
Previous to your birth did your parents live in the country, town or city?
What portion of your life has been spent in the country ( ), town ( ), city ( )?
Occupation of father previous to your birth? at present?
Health previous to your birth of father? of mother?
How many hours do you sleep ( )?
At what time of day can you study best ( )?

In addition to the above, a series of questions concerning general health are submitted, supplemented by inquiries as to personal habits of eating, drinking, smoking, exercise, work, etc.

The idea of those responsible for these tests was to secure as many data as possible concerning the individual in connection with his college record and his life history. It seems certain that the correlation of such results will give us an insight into the elements that make for success in daily life, as well as point the way to scientific progress. In the following research only a part of these data has been worked over. The standpoint of the writer being largely psychological, more time has been given to the tests made in the psychological laboratory than to those of the gymnasium or class-room.

CLASS AND SEX DIFFERENCES.

Besides some 250 freshmen and a small number of seniors in Columbia College, we have the results for a number of young women in Barnard College. This enables us to compare the sexes and to estimate the changes among men during their college life. In Table I. are such results as can be easily reduced to averages. Here are the totals for freshmen, seniors and women. The probable distribution of cases around each average is computed according to the formula

$$p = 0.6745\sqrt{\frac{\Sigma v^2}{n-1}}$$

In length of head, for example, $p = 4.3$, and 194.0 ± 4.3 is understood as defining the limits between which approximately 50% of the cases will fall. p expresses the probable variation in the group, and the limits of the group may be estimated as average $\pm 5p$.

In estimating the certainty of a difference between two averages the rule is to proceed upon the principle that the combined error of these averages is equal to the square root of the sum of the squares of their respective errors. The relation of the actual difference to this combined error furnishes the basis for estimating the certainty of the result. In Table I. this is expressed in terms of unity under the heading C.[1]

For the sake of clearness the important class and sex differences are shown graphically in Fig. 1. Here the averages for freshmen are taken as a base and the direction and amount of difference plotted accordingly. In reading these curves, *above* means superiority in size, strength, sensitiveness, accuracy, quickness, quality and quantity. The amount of difference is expressed in terms of p and indicated on the ordi-

[1] For methods of calculation and deduction of formulas see Merriman, A Text-Book of Least Squares; Westergaard, Die Grundzüge der Theorie der Statistik, etc.

TABLE I.

	Sex Differences.							Class Differences.					
	Men.			Women.			C	Freshmen.			Seniors.		C
	Cases.	Av.	$\pm p$	Cases.	Av.	$\pm p$		Cases.	Av.	$\pm p$	Av.	$\pm p$	
Head—length (mm.)	267	194.0	4.3	42	181.5	5.8	0.99	34	193.1	4.0	195.3	4.4	0.86
—breadth "	267	152.9	3.8	42	148.1	3.1	.99	34	153.5	2.6	155.5	3.4	.93
Strength of hand—right (kg.)	210	36.3	4.9	42	20.0	3.6	.99	35	38.0	5.3	41.2	5.9	.74
—left "	210	33.5	4.7	42	17.5	3.5	.99	35	34.4	4.9	37.0	5.0	.77
Fatigue (%)	110	65.1	27.0	39	65.0	25.1	.05	......	......	......	......	......	......
Perception of weight—error (g.)	123	144.1	51.7	35	140.	60.3	.16	......	......	......	......	......	......
Pain—right hand (kg.)	266	5.9	2.4	42	2.4	1.2	.99	35	6.4	2.8	5.7	1.4	.65
left hand (kg.)	266	5.6	2.2	42	2.4	1.1	.99	35	5.6	2.4	5.5	1.5	.11
Sensation areas—av. correct in 5	262	3.7	0.9	42	3.8	0.8	.41	33	3.4	0.9	2.8	0.7	.97
Pitch—error on 75 cm. string	265	7.2	4.7	42	5.3	2.9	.99	33	7.6	4.2	3.7	2.5	.99
Perception of size—error (mm.)	243	2.4	2.0	34	3.4	2.1	.92	35	2.1	1.2	2.7	1.2	.86
Estimation of 10 sec.—error (sec.)	127	1.7	0.8	......	......	......	......	31	1.8	0.6	1.3	0.6	.98
Reaction-time (σ)	256	158.7	18.9	42	183.9	27.7	.99	47	164.0	16.0	157.2	20.0	.80
Marking out A's—time (sec.)	270	100.2	12.2	42	91.2	8.1	.99	35	105.4	12.8	88.9	10.3	.99
—A's omitted	270	2.2	1.6	42	3.0	1.9	.92	35	4.7	2.0	1.6	0.7	.99
Naming 100 colors—time (sec.)	123	84.8	14.2	40	65.1	8.3	.99	33	......	......	79.0	18.6	(.61)
Movement-time (sec.)	59	34.1	4.3	23	36.0	4.0	.82	33	......	......	28.6	2.5	(.99)
Accuracy of movement—time (sec.)	58	49.0	7.0	......	......	......	......	......	......	......	......	......	......
—error (mm.)	58	0.8	0.3	......	......	......	......	......	......	......	......	......	......
Association-time (sec.)	171	55.4	22.9	39	74.8	26.0	.99	32	53.8	23.1	47.5	21.0	.54
Auditory memory—numerals recalled	266	7.6	0.4	42	7.3	0.5	.97	34	7.6	0.5	7.6	0.4	.00
Visual " "	142	6.9	0.5	42	7.3	0.4	.99	33	......	......	7.0	0.4	(.58)
Logical memory (%)	101	44.5	11.1	37	48.2	13.2	.65	33	......	......	58.5	12.0	(.99)
Retrospective memory—error (mm.)	256	4.5	2.6	......	......	......	......	31	6.3	2.7	5.4	3.2	.58

nates in a scale of 10 p. In this way we are able to express the relative differences for all averages and so have a common unit of measure. In Table I. these differences have been translated into probabilities and we have assumed that when the probability of a difference is 0.90, or 2.5 p, it may be taken as a difference of considerable certainty. This limit we have indicated

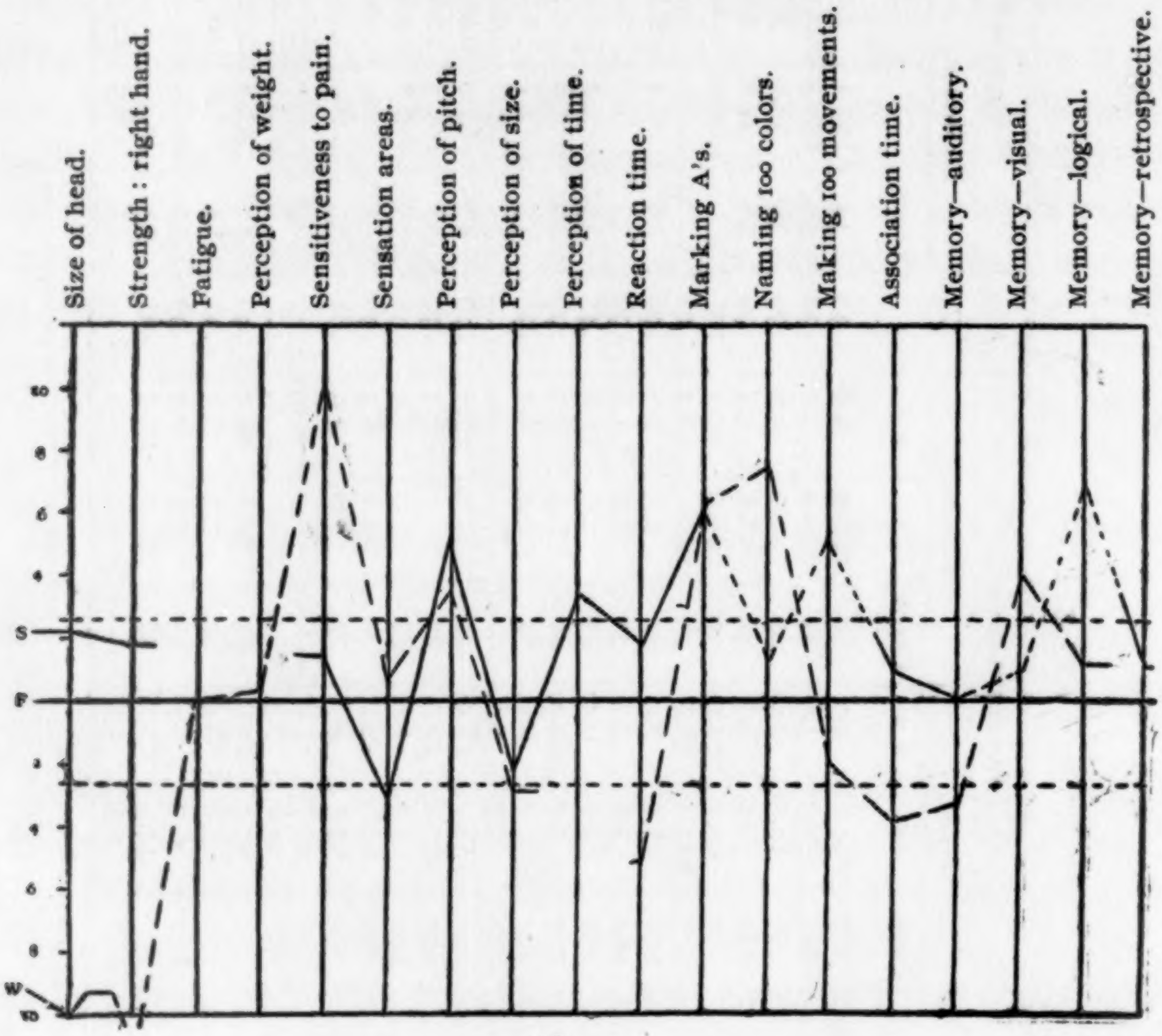

FIG. 1.

by the parallel broken lines. A difference of 5 p has a probability of 0.999 which may be regarded as sufficiently certain for all purposes. In Table I. a certainty of 0.82 may be regarded as indicating a probable difference and 0.50 as equivalent to no difference. In reading the curves the probability of all differences should be estimated according to the above. Readers unfamiliar with such methods of treating data should bear in mind that this is only a way of estimating the allowance that must be made for accidental variation.

According to this we may say that the women are approxi-

mately equal to men in fatigue, perception of weight, sensation areas and logical memory; they surpass the men in sensitiveness to pain, perception of pitch, marking out A's, naming colors and in visual memory; they are inferior to them in size of head, strength of hand, reaction-time and association-time; they are also probably inferior in movement-time, perception of size and auditory memory.

With the seniors and freshmen we have the record of each individual for comparison, and it is these differences that are plotted, except where the line is broken. So far as we have data, the seniors surpass themselves as freshmen in width of head, perception of pitch, estimation of time, marking the A's and are inferior in sensation areas: they are also probably superior in length of head, strength of hand and reaction-time and inferior in the perception of size. As compared with the class results of freshmen they are superior in movement-time and logical memory. It will be observed that in case of pain, association-time, and visual and retrospective memory they are actually above the freshmen averages. The other tests not represented in the plotted differences may pass without comment.

In addition to the results given in Table I. the following seem of sufficient interest to deserve mention.

One and a half per cent. of the freshmen were color blind, but no cases were found among the women. In color preference 42% of the men chose blue and 42% of the women red, these being the preferred colors. The total distribution of likes and dislikes is as follows:

	Freshmen.		Women.	
	Like %.	Dislike %.	Like %.	Dislike %.
Red	22	7	42	8
Orange	5	25	8	31
Yellow.	2	32	5	8
Green	7	15	9	21
Blue	42	12	9	23
Violet	19	8	19	9
White	3	1	8	0

Of the 28 seniors 20 showed the same preference as when freshmen and of these 8 were for blue. All the senior prefer-

ences fell upon red, blue and violet except two. In brief, the tendency seemed to be a shifting toward the violet end of the spectrum as age increased.

In eyesight there is no certain difference between the sexes unless it be in the variability. The seniors as a class are neither better nor worse than in their freshmen year, though some individuals have lost and some improved. The distribution of cases was as follows:

EYESIGHT TESTS.

Cm.	Freshmen.		Women.	
	R %.	L %.	R %.	L %.
85	0	2	0	0
72	15	14	10	5
61	26	26	22	21
52	26	34	40	40
44	13	12	22	19
37	12	7	0	5
31	7	0	2	4
26	0	2	2	2
22	0	0	0	0
19	1	0	0	2
16	0	3	2	2

According to cephalic index freshmen and women were distributed as follows:

Index	70–74	75–79	80–84	85–89
Freshmen	11%	50%	33%	6%
Women	6	50	39	5

So far as these results go it can not be said that there is a real difference between the sexes.

Ellis[1] has brought together certain results which lead to the conclusion that to reach the maximum at the first trial with the hand dynamometer is a sign of weakness and that this is further evidenced by the predominance of such relation among women. The results of these tests furnish data bearing directly upon this point. Among the freshmen in Columbia College 61% reached the maximum with the right hand at the first trial in the test and 63% with the left. The distribution is apparently the same for both hands, but the student who reaches the maximum record on the first trial with the right hand is no more likely to do

[1] Man and Woman, p. 151.

so with the left hand than those who do not. Taking those who reach the maximum on the second trial, we find them not among the strong men but equally distributed among all grades. So it appears that this variation is independent of absolute strength. However, we find that 80% of the women reach the maximum at the first trial for the right hand and 66% for the left. This may be an accidental difference, but should it prove to be a secondary sexual characteristic the explanation cited above is an unwarranted assumption.

The association test furnishes a large number of associations for each word. In the senior and freshman tests we have two sets of associations with an interval sufficient to eliminate memory interference. Of these 34 cases we find the number of identical associations in the two tests as follows: for house, 9; tree, 8; child, 10; time, 4; art, 9; London, 9; Napoleon, 7; red, 10; enough, 5. On the whole this is more of an agreement than we should have supposed, and indicates a persistence of early associations.

A classification of associations shows that 25% of the Barnard tests give particular associations, as Barnard House, 49th St., etc., while the men give but 6% such associations. From reading over all the lists we get the impression that the women excelled the men in particular and personal associations. As to the kinds of association, no particular difference between the sexes appears. The comparative results of freshmen and seniors give no marked changes in the character of associations.

The answers to the questionnaire on imagery show no important sex differences. Questions as to the nature of our mental images are always difficult to answer and are probably especially so for people not psychologically inclined. The seniors who answered the questions after an interval of four years show great variation, the greatest uniformity occurring in the two first questions of each group in the blank. Eighty-five per cent. of the freshmen and 88% of the women found the outlines distinct and sharp, 68 and 69% could image a face better than a voice. With respect to real hallucinations, 20% of the freshmen reported as against 40% of the women. About 65% of such experiences were auditory. More than half the freshmen reporting experienced

repeated persistent hallucinations of being called by name and of these about half were usually recognizable as "mother's voice." As this is a subject of general interest we give a few typical reports:

"Besides having heard the voice of mother when young while she was some 1500 miles away, I now and then hear scraps of conversation (which I have never heard) float before my ears and also voices of friends call me by name. This, however, is seldom, and I think occurs only when I am greatly fatigued in mind or body."

In cases of visual experience the image is usually that of an unknown person, but one often saw faces belonging to persons known to be dead. Here is the one striking case of a familiar figure:

"One night I awakened from a deep sleep and saw my father standing at the foot of the bed. The image was perfectly clear and appeared to my mother a few nights afterward."

From the senior reports we may see what changes have come during four years. Ten seniors reported cases in the freshmen tests. In the senior tests no new cases were reported, one reported no recurrence of the experience, four reported persistence of the same experience and five recurrence with modifications. The last show a slow development of the experience. One reported that when at a preparatory school he sometimes heard his mother's voice when a piano was played. Now when playing the piano away from home usually in the midst of a selection he hears his mother call in a loud voice, always causing a break at that point. This is a constant experience and becoming more and more frequent.

Although the number of cases is small the tendency seems clear; that the majority of such experiences are auditory and of these the hearing of one's own name in a strange or familiar voice is most common; that, whatever the nature of the experience, it is usually persistent and develops along its own beaten paths—those who had visual experiences have them still, etc. All seem to have their origin before college life begins. However, these are to be taken as suggestions rather than definite conclusions.

In the memory tests for numerals a decided interference effect is often observed in that the second series is confused with

the first. This is especially true of seniors and mature students, with whom the probability of the first trial being correct is much greater than in case of the succeeding. In retrospective memory the error may be measured in two ways, according to the deviation from the standard line or according to the deviation from the line actually drawn in the size test. The results show that the average error for the latter is about half as large as for the former. This difference is doubtless due to a constant error in the individual.

The miscellaneous data found in the various blanks do not differ materially from those given in the preliminary report. As a brief summary, we find freshmen in Columbia College about 18 years of age, about 90% are regarded as having brown hair, their eyes are about equally divided between gray, blue and brown, 70% have straight foreheads, 60% thick lips, 75% are of clear complexion and about 2% are left-handed.

Since sex and class differences are incidental to the purpose of this research, no attempt has been made to compare these results with other data published elsewhere. Difference in method of testing rendering such comparison difficult in any event, it appears best to limit this section to a statement of specific results.

It remains to consider the differences between the freshmen tested for each year. Taking length of head as a special case, we find the following:

Class.	Cases.	Av.	$\pm p$.
1900	60	193.5	4.4
1901	66	193.6	4.0
1902	82	194.4	5.0
1903	59	193.9	4.7

The maximum difference is 0.9, which has a probability of 0.54, or equivalent to an accidental difference. As to how far this is true of other tests may be inferred from the following table of probabilities similarly determined.

Reaction-time,	0.98
Marking A's,	0.77
Naming colors,	0.68
Association-time,	0.68
Logical memory,	0.50
Pitch,	0.25
Pain,	0.82
Strength of hand,	0.16

Thus, with the exception of reaction-time, we find all the maximum differences between the freshmen of the various college classes within the limits of accidental variation. This exception, however, is not very important, since it is found to be a difference of but 3.5 *p*.

The import of this result is that with respect to these tests freshmen entering Columbia College from year to year are similar, or represent a type. This is also borne out by the regularity of the distribution for the various tests.

THE CORRELATION OF RESULTS.

The most important consideration in a series of tests is their correlation. It is desirable to know to what extent ability in one case assures us of ability in another. This is the real question in all tests. But before going into the subject proper some attention must be given to method. Let us take for example the stature and weight of students, as a case in which we should expect some correlation. In Table II. the cases are grouped in differences of 5 kg. and 5 cm. and so arranged that the weight

TABLE II.

Stature in cm.	Weight in kg.								Total.
	45–50	51–55	56–60	61–65	66–70	71–75	76–80	81–85	
155–160	1	1		1					3
161–165	1	2	7	1	2				13
166–170	1	5	7	10	3				26
171–175		2	9	11	8	2	1		33
176–180			4	11	6	3	2	2	28
181–185			1	3	3	4	2	3	16
186–190								1	1
Total.	3	10	28	37	22	9	5	6	120
Av. Stature.	162.5	166.5	169.8	172.8	173.6	178.6	178.5	181.6	172.5

of any group can be compared with its stature. The average stature of all is 172.5 cm., weight 63.4 kg. At the foot of each column is the average stature for the corresponding weight. Here we see that those of average weight are also of average stature and that weight increases with stature. Had there been no relation between weight and stature, the averages for the columns would have been approximately the same.

These relations may be presented in graphic form as in Fig. 2. Let *AB* represent the place of average statures, if there were no correlation; *CD* their place in perfect correlation; the point of intersection, the place of the group having both average stature and average weight; then plotting the actual averages gives the line *EF*.[1]

[1] See Grammar of Science, Pearson.

In estimating the degree of correlation it is customary to proceed by the formula

$$r = \frac{\Sigma xy}{n\mu_1\mu_2},$$

in which x and y represent the variations of an individual from the respective averages of the two distributions to be correlated, μ_1 and μ_2 the mean square deviations of the distributions and r the coefficient of correlation.[1] When r equals unity correlation

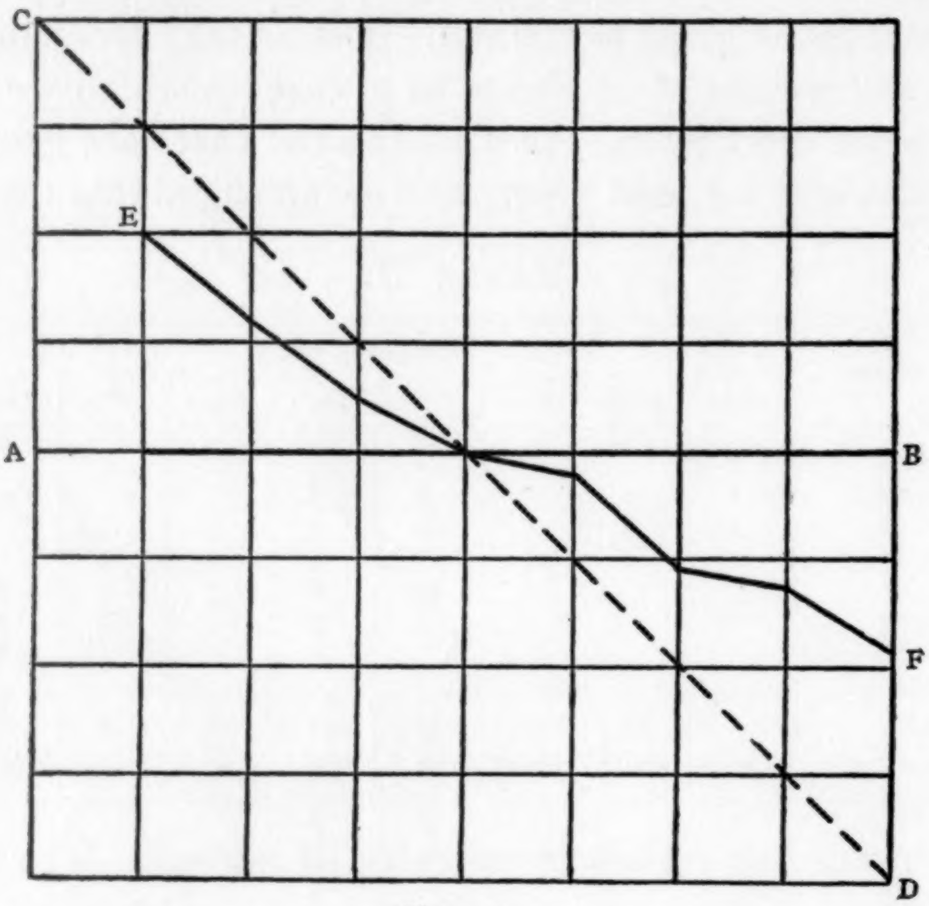

FIG. 2.

is perfect, when equal to zero the correlation is that of chance. Applying this formula to the data for stature and weight (Table II.) we find $r = 0.66$.[2] This coefficient expresses the relation between the averages for the successive columns, or arrays, in the table or the place of *EF* in Fig. 2. The probable variation of an array may be found from $p\sqrt{1 - r^2}$. Since $r = 0.66$,

[1] Natural Inheritance, Francis Galton. Mathematical Contributions to the Theory of Evolution, Karl Pearson, *Transactions of the Royal Society of London;* also Grammar of Science, Revised Ed. The Cephalic Index, Franz Boas, *American Anthropologist*, Vol. I., No. 3, N. S.

[2] No account was taken of differences in age.

this expression becomes $0.75\,p$; calculating p for all statures in the table gives 4.6 cm., from which it follows that the probable variation for an array should be 3.3 cm. Calculating the actual variation in statures for weights of 56–60, 61–65 and 66–70 we obtain 3.5, 3.7 and 3.8. The significance of $r = 0.66$ is now apparent; if we select students according to weight, their statures will fall within approximately ¾ the limits for students selected at random. In the same way it may be shown that a coefficient of $0.43 = 0.90\,p$; $0.66 = 0.75\,p$; $0.87 = 0.50\,p$; $0.97 = 0.25\,p$; etc., from which the significance of the various coefficients in the following pages may be estimated.

Since the coefficients of correlation depend upon the mean square deviations, their own probable errors can only be estimated in the same terms. According to the method employed, it appears that for the following data differences greater than 0.10 have considerable certainty.

In order to save time the writer used another form of this method which may be illustrated by an actual case. In reaction the cases may be divided roughly into four groups as

(1) $$68 + 70 + 55 + 63 = 256$$

and in the A test

(2) $$57 + 80 + 66 + 53 = 256.$$

Now, if a mere chance relation held between these tests, the 68 cases in reaction should be distributed in series 2 as

$$\frac{57}{256} + \frac{80}{256} + \frac{66}{256} + \frac{53}{256} \text{ of } 68$$

or

$$15.14 + 21.24 + 17.50 + 14.07 = 67.95.$$

Counting through, we find these 68 cases actually distributed as

$$15 + 20 + 15 + 18 = 68.$$

Taking the 63 in series 1, the chance occurrence should be

$$14.02 + 19.68 + 16.24 + 13.04 = 62.98.$$

The actual order of cases was

$$12 + 21 + 18 + 12 = 63.$$

Thus we find the correlation approximately that of chance. The method can be shortened by calculating the number of cases probably agreeing in the corresponding groups in series 1 and 2.

In the same manner for stature and weight,

$$(1)\ \text{Stature}\ 42 + 33 + 45 = 120$$
$$(2)\ \text{Weight}\ 41 + 37 + 42 = 120.$$

The 42 short men in 1 occur in 2 as

$$25 + 12 + 5 = 42.$$

By chance they should be

$$14.4 + 12.9 + 14.7 = 42.0.$$

The tall men occur

$$5 + 14 + 26 = 45.$$

By chance,

$$15.3 + 13.8 + 15.7 = 44.8.$$

Thus we find evidences of correlation in this case.

In a series of tests like the following, where there are over six hundred possible correlations to consider, a shorter method still will be useful. Correlation may be detected by simply taking account of the distribution of cases. Returning to Table II. we find in regard to stature about the same number above as below the average group. Arranging the 42 short men, the 45 tall and the 33 average men according to weight, we find them distributed as follows:

Short men,	3	8	14	12	5	0	0	0	42
Tall men,	0	0	5	14	9	7	4	6	45
Av. men,	0	2	9	11	8	2	1	0	33
Total,	3	10	28	37	22	9	5	6	120

That tall men are heavier than short men is now evident. In practice we need but find the distribution of cases with high and low values, say above $+p$ and below $-p$. A judicious use of this method makes a rough-and-ready estimate of the degree of correlation possible in any case.

The data discussed in the following pages were first arranged according to this last method; then the actual distributions were compared with the theoretical chance distributions; finally all the important tests and all giving any evidence of correlation were treated by the Pearson formula and the coefficients taken as the basis of discussion. Thus the treatment has been sufficiently thorough to give a reliable conclusion as to the general degree of correlation between such tests as are considered.

Tests of Quickness and Accuracy.

Reaction-time and marking out the A's furnish 252 cases in common. The reaction-time for each individual is the average of five to three valid reactions, observers recording five reactions. In the A test it has been proposed to lengthen the time recorded proportionally to the numbers of A's missed.[1] But this is unnecessary, since $r = -0.05$ in case of A times so adjusted and -0.07 for them regardless of errors. It appears, then, that the degree of correlation between these tests is approximately zero, or a chance relation. In other words an individual with a quick reaction-time is no more likely to be quick in marking out the A's than one with a slow reaction-time.

The other correlations for tests of quickness may be enumerated as follows:

	Cases.	r
Reaction and naming the colors	118	0.15
Reaction and association	153	0.08
Marking A's and naming the colors	159	0.21
Movement time and naming the colors	97	0.19
Movement time and reaction	90	0.14

The remaining possible correlations in this group of tests are movement time with A's and association, and naming colors and association, but when tabulated the distributions were such as to make a very low degree of correlation certain. Thus it appears that the time required for naming the colors correlates better than any test in this series, yet the coefficients are too low to be of much significance. ($p\sqrt{1-(0.20)^2} = 0.98p$). The results are given in tabulated form for reaction-time and A's (Table III.) and naming colors and A's (Table IV.).

Though these tests fail to correlate with each other, it may yet be true that the individual who happens to hold the same rank in two tests will tend to hold that rank in all and so divide our subjects into groups one of which correlates and the other not. This question may be answered in the following manner: From the value of p (Table I.) we may divide the individuals tested into four groups, approximately equal. In reaction-time and the A test there are 68 and 57 cases, respectively, in the first

[1] Cattell and Farrand, Psychological Review, Vol. III., p. 618.

TABLE III.

Reaction time.	Marking A's.								Total.
	60–69	70–79	80–89	90–99	100–109	110–119	120–129	130–139	
100				1			1		2
110			4	8	1	1	1	3	18
120	2	5	4	4	4	3	1	2	25
130	1	3	4	11	3	5	1	1	29
140	2	4	6	10	7	5	5	2	41
150	1	3	7	6	4	4	1		26
160		1	6	7	7	3	2	2	28
170		3	5	6	6	5	2		27
180	3		4	6	5			1	19
190	1		6	3	1				11
200		1	2	2	2	1	1	1	10
210	2		3	2	2				9
220			2	2	2	1			7
Total.	12	20	53	68	44	28	15	12	252
Av.	145	143	162	151	160	150	144	140	

group, or upper quarter. Of these 15 are identical, or the number conformable to chance. In the classes of 1902 and 1903 there were 30 cases in the upper quarter for naming colors;

TABLE IV.

Time for marking A's.	Time for naming colors.									Total.
	50–59	60–69	70–79	80–89	90–99	100–109	110–119	120–129	130–139	
60–69		1	3	1						5
70–79	4	3	4	1	3	1	1		1	18
80–89	6	9	11	6	3	1	1	1		38
90–99	2	10	9	5	4	1	1		2	34
100–109		3	6	7	4	4	4	4	1	33
110–119	3	6	2	3	2	1	1	2	2	22
120–129		1	2			1				4
130–140			2	1	1				1	5
Total.	15	33	39	24	17	9	8	7	7	159
Av.	90	95	95	98	97	102	99	119	115	

this test seemed to correlate in a way with reaction and the A test, yet there is but one case identical in the upper quarters of the three. Theoretically, the chance agreement is 0.4 of a case. The other quarters give corresponding results.

The cases are not sufficiently numerous to warrant giving more space to this correlation. At least the results indicate that, though we find some reasons for assuming double correlations

between certain tests, the rank of the individual in the whole series of time tests seems subject to chance. This conclusion is out of harmony with all general belief, since we usually assume that quickness is a fundamental thing characterizing the individual in every respect. But the discussion of this point must be deferred.

Having now considered tests of quickness, we may take up those whose results are generally expressed in terms of accuracy. In the size test there are three similar requirements: drawing a line, bisecting it and bisecting a right angle. Some correlation appears between the former, 0.38 (123). The errors in all cases are about equally divided between positive and negative, and those who underestimate the line are no more accurate in its bisection than those who overestimate it. Curiously enough, when we come to correlate these results with bisection of the angle, no agreement save that of chance appears.

The length of the perpendicular line involves an illusion and is not so easily correlated. On the average this line is $\frac{5}{6}$ of the actual base, only 2% drawing it longer than the base.

The tests for accuracy of movement (striking the dots) and the perception of weight (force of movement) correlate neither with each other nor with the test for size. Also, no correlations were found between these tests and the accuracy of estimating time intervals or of following a given rhythm. The A test may be considered here also, since it may be graded according to the number omitted, but this fails to correlate with the above.

It may be well to consider the relation of speed to accuracy. In case of A's they hold a slight inverse relation, $r = -0.28$. The time required for naming the colors which seemed favorable to correlation in the group of time tests does not correlate with any of the tests in this group. There is an apparent correlation between the accuracy of movement test and the time taken to do it; however, the writer regards the results of this test as uncertain, because any individual can hit all the dots if he goes slowly enough.

It may be urged that in many cases failure to correlate is due to want of precision in the tests, so we must give some consideration to the validity of the methods employed. In the A

test, for example, we have but one trial, and unless we know something of the limits of variation in repeating it we can give no particular value to the results. It is difficult to repeat this test without modification of some kind; if the same blank is used, the subject remembers where some of the A's were found and thus gains in time; if a new blank is made the conditions vary. We give here the results for two men and two women who made repetitions of the test in succession.

Men	94	90	91	91	80
	70	66	67	66	67
Women	83	74	71	75	70
	82	76	75	69	71

Two other subjects made one daily trial for six days:

62	60	53	60	57	60
70	68	66	62	59	60

Thus we see what might have been foretold. The first trial is likely to be the longest for want of practice, but it does not appear that the difference between the first and second trial is out of proportion to the succeeding, conforming generally to a practice curve. This satisfies the requirement of an efficient test.

It will be remembered that the seniors repeated this test after an interval of four years; the distribution of cases as freshmen was according to *p*, 7–10–9–9, and of these 5–7–8–5 respectively held their relative positions.

In naming the colors there should be relatively little variation between the various trials since this is essentially a reading time. Five freshmen made two trials of the test at intervals varying from one hour to three days with an average first trial of 61 seconds, second trial 58. We find here proportionally smaller differences than in case of the A test.

With movement-time the difference is relatively large, since the individual variability is about the same as that of the group. In repeated trials with five subjects the average difference between the first and second was 4 seconds, all shorter times; between second and third, 1.9 seconds. The former is approximately the variability of the group. The apparent regularity of gain in the different subjects, however, gives this test some certainty.

In case of the accuracy of movement, we have no data except the daily records of two subjects.

Time,	47	36	36	35	35	32	40
Error,	0.9	1.0	0.8	0.8	0.8	0.8	1.1
Time,	40	37	34	29	30	28	30
Error,	0.8	0.9	1.0	1.0	1.3	1.4	1.2

Here we observe some individual differences. One subject shows a tendency to sacrifice accuracy for speed, the other gains in speed with a fairly constant accuracy. The inference would be that the result of this test will vary according to the attitude the individual assumes. Unfortunately, we have too few cases for precise correlation.

In the association test we have no data for repeated trials except in case of seniors, who gave even more decided agreement in rank than in case of the A test. For this reason we are justified in placing some confidence in the result. Whether the test measures the readiness of association is doubtful, but it seems that the time is indicative of the reaction to the test and that individuality in this respect persists during college life.

When we consider reaction-times we have five trials for each individual and have to do with an average within an average, from which it follows that the precision of the test will depend upon the probable variation of the group. Taking ten cases at random from freshmen of 1900, we find the averages giving an average probable error of 7σ, and as seniors the same individuals giving an error of 12σ. The extreme case was an error of 18σ. The value of p for the group is 19σ. Thus the probable errors of the individual averages are within the limits of the group variation. In the preliminary report it was shown that the tendency was toward quicker reactions in the successive trials. Five freshmen made ten reactions under the conditions of the test and the average of the first five was 10σ longer than the average total. The repetition of some cases several days later and an extended series upon two other individuals show a much smaller difference. Here again we have a practice effect and it appears that the tendency is general. The seniors held the relative rank of their freshmen reaction tests as well as in other cases.

In case of the size test we find the distribution of errors for two individuals with a 10-cm. line, 12 daily trials of 10 lines each, 120 lines in all, in every way comparable to the group distribution for the 5-cm. line. In an extended series of trials with the 5-cm. line with three individuals the average errors were 2.5 ± 2.2; 6.0 ± 0.9; 2.0 ± 0.7. In the case of seniors we find that they do not keep their relative rank as freshmen in any sense. Though we can not say that this test is very precise, the fact that it correlates in itself seems to indicate sufficient precision to reveal correlations with other tests, if such exist. The other tests of accuracy may pass without comment.

This leads to the conclusion that all the time tests, excepting possibly the rate of movement, are sufficiently precise to bring out correlations if there were decided tendencies in that direction. In short, the failure to give decided correlations can scarcely be attributed to that source. While it is true that variation of the individual is often large, as in case of reaction-time and rate of movement, it is improbable that those above the value of p should in reality belong to the subgroup below that value. Nor does it seem that the tests of accuracy fail to satisfy the requirement of precision in this respect.

Memory Tests.

Any method devised for the gradation of memory tests must be arbitrary. This is especially true in logical memory, and renders comparison of results impossible except when graded by the same individual. In the case of numerals, taking the number correctly written is the most precise, since, if we grade by order, we must make arbitrary estimates. In this study we have graded by the number correctly written and the number in position. It will be said that grading by position is not advisable, since by the omission of one numeral alone a whole series may be counted as out of place; but such combinations were of rare occurrence in these tests, and when they did occur the case was thrown out.

There would be no reason for considering the number correctly written if it were closely correlated with the number correctly placed. But such correlation is not apparent. In the

auditory test half the freshmen (135–266) gave all the numerals correctly and one-seventh (20–142) in the visual test. The former does not give sufficient distribution for precise correlation but, if we take the number correctly placed, the cases correct in all are approximately one-half the above groups. That is to say that the chances are even that an individual who can recall a series of eight numerals will be able to give them in correct order. We can not enter into a discussion of the independence or dependence of ability to recall the individuals in a series and the ability to reproduce the series itself. But if we correlate the number correctly written in the auditory and visual tests we find with 144 cases a coefficient of 0.29 and in the case of number correctly placed 0.39. This is somewhat above a chance relation. If it were safe to put much faith in small differences, we might give some weight to the consistency of these results, but as previously noted the distribution of cases is such that correlations for number correctly written must remain uncertain.

The logical test was graded in a scale of ten and the distribution of cases was found quite regular, affording ample opportunity for correlation. For position in the auditory test a correlation of 0.04 (94 cases) results; for number correctly written 0.05. By use of the short method there seemed no evidence of a marked difference between these and the probable coefficients for visual memory. On the whole, there seems to be a decided absence of correlation between this and the numeral tests.

In the case of the line drawn from memory we find a correlation with logical memory of −0.07 (91 cases). There was no apparent correlation with the numeral test.

The test for pitch is in many respects a test of memory, and may be taken up here. As seen in Table I., the variability is such as to give a wide distribution of cases. From a tabulation of results it appeared that about two-thirds of the cases fell into a typical group with an error of 4 cm. ($\frac{4}{75}$ of the string), while the remaining cases were widely scattered beyond that limit. This makes the number of available cases small, except in case of auditory memory; yet there were no signs of correlation with the

memory tests. But this method of selecting cases seems uncertain, since a study of variation in individual cases indicates that the present test cannot be considered precise below differences of a half tone. Taking the two trials of the freshmen, we find that the chances that the second error will vary from the first within the limits of a quarter tone are 65 in 100; within the limits of a half tone, 85 in 100. With differences of a half tone the distribution of cases gives a regular asymmetrical curve, but this correlates with the other memory tests no better than the selected group.

The test for force of movement is also largely a memory test, and, though its results gave regular distributions, it failed to show any evidence of correlation with the other tests. We might expect that it would in drawing the line in the size test, but the result was −0.08.

In general, we may say that these tests fail to show much intercorrelation. The numeral tests, it is true, correlate somewhat, but that would be expected; not necessarily because of a functional relation, but because the individual is likely to mentally repeat the numerals in either test. The data are not at hand for determining the relative value of these tests or even their precision. The whole subject of memory tests needs further consideration before positive conclusions can be reached.

Since it appeared that certain tests for quickness tended to correlate, though with low coefficients, it may be well to try them in the memory tests. We find naming colors gives 0.03 (93 cases) with logical memory; reaction-time for logical memory 0.12 (96), for position in auditory memory 0.17 (112) and for position in visual memory 0.06 (104). This seems to indicate that a quick reaction on sound has more weight in memory for things heard than seen, yet the influence is too slight to be of much consequence. Also reaction and pitch gave 0.01 (100).

College Standing.

In general it appears that correlations in any of the foregoing tests are not of a degree sufficient for practical purposes. We do not learn much of an individual by any one or any group of them. It appears that we are dealing here with special and

quite independent abilities and that the importance attributed to such measurements of elementary processes by many investigators is not justified in this case. However, it seems probable that a basis for correlation will be found somewhere, and it may turn out that though these tests do not show much intercorrelation they may individually correlate with ability in the more complex tasks of life. Research in this direction is obviously difficult for want of adequate standards and satisfactory data, and the only attempt we have made is in respect to college standing.

Students in Columbia College are marked in a scale of five for each semester. The standing for the year is estimated by the sum of the products of the grades and the number of course hours, divided by the total number of such hours, or the average grade per course hour. The distribution of these averages for 240 cases was as follows, it being understood that 1.0 represents the highest possible grade.

Grade,	1.0–1.2	1.3–1.6	1.7–2.0	2.1–2.4	2.5–2.8	2.9–3.2	3.3–3.6
Cases (%),	2.1	7.5	13.7	15.4	19.1	15.8	13.3
Grade,	3.7–4.0	4.1–4.4	4.5–4.8	4.9–5.2	Total		
Cases (%),	6.2	4.5	1.6	0.4	99.6%		

This gives a distribution sufficiently regular for correlation. It may be said in passing that the character of this distribution indicates a high degree of precision in the method of grading.

The following correlations have been calculated for the average standing:

Reaction-time,	227 cases.	— 0.02
Marking A's,	242 "	— 0.09
Association,	160 "	+ 0.08
Naming colors,	112 "	+ 0.02
Logical memory,	86 "	+ 0.19
Auditory memory (position),	121 "	+ 0.16

An application of the short method to the other mental tests gave no hope for correlation.

Here we are face to face with another cold fact: the tests of quickness seem to hold a chance relation to class standing, and ability to do well in the memory tests has but little significance. But it may be well to examine the relative standing in the different courses.

The following correlations were calculated:

Latin and mathematics,	228 cases.	0.58
Mathematics and rhetoric,	222	0.51
Latin and rhetoric,	223	0.55
French and rhetoric,	122	0.30
German and rhetoric,	132	0.61
German and mathematics,	115	0.52
Latin and French,	130	0.60
Latin and German,	129	0.61
Latin and Greek,	121	0.75
Gymnasium and average grade,	119	0.53

Here we find a higher degree of correlation than heretofore. The languages have a correlation of 0.60 to 0.75, or a reduction of variability approximating ¼. In other cases we have a reduction of about $\frac{1}{5}$. The exceptionally low correlation for French and rhetoric is rather puzzling, but is probably due to some accidental cause. On the other hand, the high degree of correlation between Latin and Greek is according to expectation. It is interesting to note that the degree of correlation here is about the same as for stature and weight. From what has gone before it is improbable that a high degree of correlation will be found between particular courses and the separate tests. For example, with logical memory and mathematics the coefficient is 0.11 and with Latin 0.22, no significant variation from 0.19, the correlation for the average class standing.

Whatever it is that makes for correlation in class standing seems to hold generally for all courses. The gymnasium grade, which is based chiefly upon faithfulness in attendance, correlates with the average class standing to about the same degree as one course with another. We have not carried this correlation out to its full possibility, as that would take us too far afield. Yet this serves as a suggestion as to how this method of correlation may aid in the solution of a very important phase of the test question. We might extend it so as to embrace all courses offered, work out the degree of correlation for the different college years or semesters, correlate the successive marks for the students of one department or of a single instructor, etc., all bearing upon interesting and live questions.

Physical Tests.

The strength and fatigue of hand have been objects of much attention for many years, but notwithstanding the mass of published data and the great diversity of mechanical appliances devised, very little advance has been made. Yet it is a general belief that specific individual differences observed in such tests are correlates of mental and physical differences not far removed from the fundamental. It has been claimed that strength of hand is a correlate of mental ability, that civilized men are stronger than uncivilized, and professional men than laborers. In these tests we find no correlation between class standing and strength of hand, $r = -0.08$. For reaction-time, marking out A's and naming the colors there is some correlation but with very small coefficients. It might be inferred from this that general physical vigor will have some weight in mental tests such as we have in this series, but this is by no means certain, for we make this inference upon the assumption that strength of hand is an index of general physical vigor. The fatigue test employs the muscles of thumb and forefinger only. If we take the total reading of the ergometer for the fifty efforts we find little correlation with the strength of hand as measured by the dynamometer. Taking the reading for the first ten, so as to avoid as much of the variation due to fatigue as possible, we still have a correlation of but 0.19. On the other hand, there is a close correlation between the two hands in the dynamometer test and also between the first and second trials with the same hand. This shows the difficulty of comparing the results obtained with different instruments. It will be said that we are attempting to correlate a fatigue effect with an unfatigued one; this objection is not of much weight, since the total reading of the first ten efforts in the fatigue test correlates no better with the second trials on the dynamometer than with the first or even the average of the two, which is itself modified by fatigue. Anyway, it does imply that there is little correlation between strength and fatigue tests when made on different muscles.

When we come to the fatigue test itself the question as to

how fatigue shall be estimated arises. If we take the total readings of the ergometer, we have no means of knowing to what degree strength was depleted by the test, as will be seen from the following cases:

	1st.	2d.	3d.	4th.	5th.	Total.
1.	45	40	43	42	40	210
2.	100	80	70	70	60	380
3.	70	65	75	65	67	342
4.	70	55	40	35	40	240

Cases 1 and 3 show very little change in the successive readings, while 2 and 4 show a decided decrease, or, according to the interpretation, the degree of fatigue is much less for 1 and 3. When we correlate the total readings of 144 cases with the first readings we find $r = 0.91$. This means that if we estimate by the total reading ten trials is sufficient, and then we should expect to find the total for the first ten closely correlated with the reading for the first trial; thus we fall back to a mere dynamometer test and fatigue drops out of the question.

The difficulty of devising a method by which the relative loss of ability to press on the ergometer may be estimated becomes apparent when we attempt to rank the four cases above. If we estimate according to the variation from the first reading the rank is 3, 1, 4, 2; according to the variability from the average 1, 3, 2, 4, etc. The result wished for in a fatigue test is to ascertain the endurance of individuals under like conditions, and superiority in this respect is assumed to be the degree of approximation to the initial effort. For this reason we have estimated the degree of fatigue as follows: Divide the total reading by five times the first reading. This is on the assumption that without fatigue five times the first reading will approximate the total reading, the difference between this and the actual reading expressing the degree of fatigue. Correlating these fatigue values with class standing we find $r = +0.23$, about the same degree as found for naming the colors and marking out the A's, standing in Latin and the logical memory test, and length of head and logical memory. The short method gave no indication of correlation between degree of fatigue and the other mental tests. However, the correlation between the reading for the first ten efforts and the degree of

fatigue is −0.34; that is to say, that the tendency is toward a compensation. The most natural explanation of this is that some began the test leisurely and concealed the signs of fatigue by a gradual increase of effort. Upon this assumption the above coefficient is meaningless. The writer is keenly aware of the arbitrariness of the method by which the degree of fatigue was estimated; but, if this method fails, there appears no other way free from the same objections. The enthusiastic designers of ergographs and dynamometers may say that the fault is in the instrument and that by employing their models results will be attained. The writer has used, so far as he is aware, all the types of ergograph in vogue and made a comparative study of them in the laboratory of Columbia University, with the result that when weights were used all gave similar results and likewise when springs were used other similar results. So it seems reasonable to assume that, no matter what the instrument, the result will be the same unless different muscles have different weight in correlation. In a recent study of the question[1] it was shown that certain peculiarities of the muscle in contraction, when reacting against certain peculiarities of a spring, gave results difficult to interpret and exceedingly variable. It thus becomes probable that the foregoing correlation between estimated fatigue and strength is in part an expression of the relation between the muscles and the spring of the ergometer. Thus a problem presents itself that is beyond the limits of this paper. The fatigue test needs an extended comparative study on a large number of individuals.

Since we are mostly concerned with the tests predominantly mental, those of a physical character may be passed over hastily. In case of head measurements we find a correlation of 0.21 (99 cases) for logical memory and length of head, but for breadth of head −0.05. By the short method no appreciable correlations were found between size of head and the other memory tests, for reaction-time, or naming the colors. As previously stated, some correlations appeared between length of head and vital capacity, vital capacity and stature, stature

[1] S. I. Franz. Methods of Estimating the Force of Voluntary Muscular Contraction, *American Journal of Physiology*, Vol. 6, No. 7.

and weight, and between each of these and strength of hand, but the number of cases available will not justify an attempt to determine their relative or combined weights in case of logical memory.

The correlation for memory and length of head seems very unlikely, but it may be of interest to add that by dividing the freshmen into four approximately equal groups, seven men were found to hold the same rank in all the memory tests. The corresponding head measurements were

Rank in Memory.	Length of Head.	Breadth of Head.
Good.	201	152
	202	150
Above Medium.	198	158
	194	158
Below Medium.	198	148
Poor.	190	147
	183	148

These results are rather striking and probably accidental. The writer refrains from drawing any conclusions whatever, leaving the interpretation to the reader. Tracing these seven freshmen through the other tests and college standing gave no corresponding results.

In conclusion, a few of the remaining tests may be noted. No evidences of correlation were found for pain or sensation areas. In case of color-blindness there was decidedly slowness in naming the colors. With eyesight we find a correlation with the number of omissions in the A test; that is, weak eyes result in inaccuracy, but the reverse is not true. Most of those reporting auditory hallucinations were classed as above normal in hearing ability.

Before going on to a general discussion of the subject it must be noted that we have not considered differences in age. It is possible for a close correlation with the age of individuals to neutralize a real correlation. Though we have not sufficient cases to treat each age separately, we can anticipate the result by correlating the various tests with the ages. We find the distribution of freshmen, expressed in per cents., to be as follows :

Age	15	16	17	18	19	20	21	22	23	24	25+
%	1	8	20	24+	26+	11+	4	2+	1	1+	1+

In the case of logical memory there was no correlation, for naming the colors there was a slight inverse one and for marking the A's a slight direct one. Thus it appears that the influence of age is small in the mental tests. Its influence in the physical tests we have not time to estimate, but it is probably greater than in the mental tests.

Also, it may be well to consider the effect of combining the results of the various tests in such a way as to give each individual an average standing, comparable to class standing in college. If it is true that such tests measure little more than special abilities, such an average should possess some significance. The calculation of such averages is by no means easy, because of arbitrariness of method. The writer used several methods and obtained satisfactory distributions of cases, but found no appreciable correlations between them and college standing. This failure may be due to faulty methods, but the probabilities favor its being the true result. Yet we should want a considerable number of separate tests before the status of the individual could be determined with much accuracy.

This does not exhaust the list of possible correlations for these tests, but represents all that could be accomplished in the time allotted. The aim has been to present in detail those of most interest to the psychologist, leaving the others to the future.

DISCUSSION OF RESULTS.

In summing up the foregoing it appears that all the tests in this series have little interdependence. The idea of correlation implies some structural or functional relation. It is generally assumed that such relation in mental phenomena is functional rather than structural, and the conception of the analysis of mental processes and the search for the fundamental elements of the same imply a close correlation between simple processes. The direction of thought is that such ideas as quickness, accuracy, regularity, etc., represent some fundamental attribute of the individual that characterizes all his acts, and hence, tests giving results in the same category must correlate. On the other hand, when it is said that correlation does not exist, it is understood that the relations are those of chance. In other words, in a given number of tests we can approximate the number of individuals quick, medium and slow in all, as well as the number for any combination of positions in the various distributions. Since this conforms to the general result of the mental tests of this series we must look for the practical value of such tests in an extended series; the superior individuals in all-around quickness, for example, are those few who happen to be quick in the majority of such tests. Thus it becomes evident that the outcome of this research raises questions which throw us back into one of the great problems of psychology, viz., What constitutes mental ability? The significance of this question becomes apparent when we consider its relation to educational practice alone. It is plain that if we accept the conclusions of this research as final, an individual must be regarded as the algebraic sum of a vast array of small abilities of almost equal probability, the resulting combination conforming to the laws of chance. But a number of objections will be raised as to the validity of such assumption. It will, doubtless, be said that such a series of tests cannot give reliable results because the observers will vary the method and the order of presenta-

tion; that practice will so modify the result that we are not sure of any one's rank; that the estimated class standing of students has no value in ranking them as to mental ability and that the differences in grades are unequal and mean nothing; that other observers have found correlations between similar tests; that the result is contradictory to all experience; that such tests are on a different level in relation to the tasks of daily life. As these are the most probable objections not previously considered they will be taken up in turn.

In such a series of tests it is necessary to have many observers. These tests were made by the officers and graduate students in the department of psychology and anthropology. None of these can be regarded as unskilled in the strict sense of the word, yet they will vary greatly in experience and interest. Each year new students are carefully instructed in the methods of procedure and taken through the tests, both as subject and observer. Still, we must consider the equation of the observer as a possible disturbing factor. There are many tests in which it seems safe to assume that any of the observers should possess all the precision necessary, such as perception of size, strength of hand, eyesight, fatigue, pitch, force of movement, rhythm, marking out A's, naming colors and movement time. In reaction, head measurements, sensation-areas, pain, and possibly memory and association, the accuracy, experience and skill of the observers may be of considerable consequence. The only way of approaching this question at present is by a comparison of the results recorded by the different observers. It so happens that the instructors and three or four of the best trained students in the department tested each year approximately three-fourths of the freshmen, the remaining fourth being tested by many observers, presumably of much less skill, each taking from two to eight men during the test period (November and December). Thus the records may be divided into groups convenient for study.

In case of marking A's, naming colors and logical memory the averages for each group of observers show differences so small that when treated by the method employed in Table I. they indicate so low a degree of certainty that they can not be

DISCUSSION OF RESULTS.

In summing up the foregoing it appears that all the tests in this series have little interdependence. The idea of correlation implies some structural or functional relation. It is generally assumed that such relation in mental phenomena is functional rather than structural, and the conception of the analysis of mental processes and the search for the fundamental elements of the same imply a close correlation between simple processes. The direction of thought is that such ideas as quickness, accuracy, regularity, etc., represent some fundamental attribute of the individual that characterizes all his acts, and hence, tests giving results in the same category must correlate. On the other hand, when it is said that correlation does not exist, it is understood that the relations are those of chance. In other words, in a given number of tests we can approximate the number of individuals quick, medium and slow in all, as well as the number for any combination of positions in the various distributions. Since this conforms to the general result of the mental tests of this series we must look for the practical value of such tests in an extended series; the superior individuals in all-around quickness, for example, are those few who happen to be quick in the majority of such tests. Thus it becomes evident that the outcome of this research raises questions which throw us back into one of the great problems of psychology, viz., What constitutes mental ability? The significance of this question becomes apparent when we consider its relation to educational practice alone. It is plain that if we accept the conclusions of this research as final, an individual must be regarded as the algebraic sum of a vast array of small abilities of almost equal probability, the resulting combination conforming to the laws of chance. But a number of objections will be raised as to the validity of such assumption. It will, doubtless, be said that such a series of tests cannot give reliable results because the observers will vary the method and the order of presenta-

tion; that practice will so modify the result that we are not sure of any one's rank; that the estimated class standing of students has no value in ranking them as to mental ability and that the differences in grades are unequal and mean nothing; that other observers have found correlations between similar tests; that the result is contradictory to all experience; that such tests are on a different level in relation to the tasks of daily life. As these are the most probable objections not previously considered they will be taken up in turn.

In such a series of tests it is necessary to have many observers. These tests were made by the officers and graduate students in the department of psychology and anthropology. None of these can be regarded as unskilled in the strict sense of the word, yet they will vary greatly in experience and interest. Each year new students are carefully instructed in the methods of procedure and taken through the tests, both as subject and observer. Still, we must consider the equation of the observer as a possible disturbing factor. There are many tests in which it seems safe to assume that any of the observers should possess all the precision necessary, such as perception of size, strength of hand, eyesight, fatigue, pitch, force of movement, rhythm, marking out A's, naming colors and movement time. In reaction, head measurements, sensation-areas, pain, and possibly memory and association, the accuracy, experience and skill of the observers may be of considerable consequence. The only way of approaching this question at present is by a comparison of the results recorded by the different observers. It so happens that the instructors and three or four of the best trained students in the department tested each year approximately three-fourths of the freshmen, the remaining fourth being tested by many observers, presumably of much less skill, each taking from two to eight men during the test period (November and December). Thus the records may be divided into groups convenient for study.

In case of marking A's, naming colors and logical memory the averages for each group of observers show differences so small that when treated by the method employed in Table I. they indicate so low a degree of certainty that they can not be

regarded as possessing any significance whatever. While the number of cases falling to any one individual is small (12 to 30), making it difficult to come to a definite conclusion, yet there seems no good reason to doubt the result. An error of a second or two in starting, stopping or reading the watch can be of little weight in a test where the probable variation of the group ranges from 12 to 15 seconds and the tasks are so easily understood that all observers could give all the directions desirable.

In reaction-time more skill and care are necessary on the part of the observers. Some experience is required to start the chronoscope properly and take the reading accurately. Taking two instructors in psychology as the expert, we find for 29 and 28 cases, respectively, average times of 147 ± 16.5 and 146 ± 19.4; the probability of the difference from the total average is expressed by 11 ± 3.4 and 12 ± 3.9, which is not a very certain difference. Taking a group of 11 observers with little experience, who represent a total of 28 cases, the average is 159 ± 16.7, almost exactly the total average. Taking two special students of psychology recording the times of 32 cases, the average is 169 ± 15.1; certainty of this difference from the total average expressed by 11 ± 3.2. It appears that so far as these results go the differences between observers count for little, as they vary quite uniformly from the total average. The difference between the instructors and the skilled students, however, is 22 ± 3.9, which reaches the limits of considerable certainty. Yet this probably represents no more than an extreme variation in subjects, since for a second group of times by two other skilled students the difference was 8 ± 3.7.

In conclusion it appears that though there is some variation in results the differences between observers can not be very great. The fact that the expert record shorter times is doubtless a sign of a real difference in precision. But to settle this question definitely we should want either a large number of cases for each observer or a long series upon the same subjects by the different observers. After working over these results the writer is of the opinion that the differences between these observers are not sufficient to give the appearance of a low degree of correlation where there is in reality a high one.

In head measurements some significant individual differences may be expected. The writer had the same head measured independently by different observers with the following result:

Observer.	Length.	Breadth.
1	194	162
2	193	155
3	195	158
4	194	159
5	195	161
6	194	163
7	192	157
8	192	158
9	192	158

Of these 1 to 6 are individuals of little experience but having acted as observers in the tests; 7 is an instrument-maker; 8 a designer and draughtsman; 9 an anthropologist of note. The correct dimensions of this head are undoubtedly the result of 7, 8 and 9. The maximum error for length of head is +3 mm., breadth +5 mm. From Table I. we see that the most probable variation of the group is 4 mm., hence these measurements are not very accurate.

There were five trained observers measuring a total of 78 heads; fourteen presumably unskilled measuring 41; and special students of the department measuring 148. The averages and mean square deviations were:

Skilled 193.8 ± 5.8
Unskilled 194.1 ± 7.2
Medium 193.9 ± 6.5
Total Av. 194.0 ± 6.7

Thus, though the averages vary little, the mean square deviations are quite different. The general rule is that the mean square deviation is inversely proportional to the square root of the number of cases, from which it follows that the skilled observers should show a deviation of 12.2. Assuming that the skilled observers give us the most accurate series, the unskilled should show a deviation of 7.9 and the medium expert 4.1. It thus becomes probable that the series is distorted by errors of measurement which are so distributed as to give equal averages but render correlation somewhat uncertain. The fact that no important correlation appeared for dimensions of head makes

it unnecessary to go into a detailed discussion of errors or to attempt to correlate the measurements of the most skillful. The breadth of head being subject to the same errors, no attention need be given it here.

The remaining tests gave differences below the limit of certainty, which means that the observer is not a disturbing factor here. A very obvious way of estimating precision would be the comparison of single observers in a given year, but thus the cases become few and the probable range of differences so great that the accepted methods of estimating their certainty give no results.

As previously stated, the subjects did not take the tests in the same order, but the series was divided into convenient groups so that several men could be taken through at once without duplicating apparatus. In few cases was the regular order of the group varied, so that the tests of each group may be considered as following a definite order. We have, then, a variation in group order; but if the results fail to correlate in other respects, the chief cause is not to be sought here, as there is no evidence that practice or fatigue correlate in such tests.

Regarding the question of practice, it is evident that, if all gained uniformly, there would be no change in rank, and hence no change in correlation; but to assert that by practice the rank of individuals will be so shifted that decided correlations will appear where none existed in case of the initial trial seems to go beyond the limits of probability. Unfortunately, we have no correlatable data for practice. The nearest approach we can make is in the case of reaction times. Correlating the first reaction with the fifth gives 0.31 (136 cases); the fourth with the fifth 0.55. The significance of these coefficients is that, while there is quite a shift in the rank of individuals in the five reactions, each has begun to find his proper level in the fourth and fifth. The effect of this change upon correlations may be inferred from the case of naming the colors: for the first reaction 0.12 (115); fourth reaction 0.20 (112); fifth reaction 0.21 (119). Those who put much faith in the element of practice as the determining factor in correlation may find these coefficients assuring, since the correlation with the average time was 0.15 (118). However, the small variation of these coefficients seems

to indicate considerable precision in correlations for the average times. In an extended study of practice reactions[1] for five subjects the relative rank was practically the same on the last day as on the first. Should this prove generally true, the coefficients of correlation with the average times in this study may be taken as the most probable result to be expected in practiced subjects. On the whole, it seems probable that practice is only one factor among many in determining correlations and so of no special weight.

The correlation for class standing being incidental to this research, no lengthy consideration of the validity of such grading need be taken up here. Whatever may be the shortcomings of the system, it is the criterion of selection from college entrance examinations to graduation, and taking the grades themselves as the point of departure the results of the previous correlations show what is accomplished by this system. In other words, it expresses the probability that a given mark in one subject will be accompanied by a corresponding mark in another. The question as to the validity of the standard and the error in its application presents a different problem, though more fundamental. The consideration of this question alone would require extensive research. The uniformity of correlation between the grades for the different courses leaves no basis for a denial of correlation, but individual assumptions as to the precision of the method in the exact statement of differences in rank will lead to unlike interpretations. The only point aimed at in this research was to show that, while the marks of students correlate with each other to a considerable degree, they show little tendency to do so with the mental tests of the psychologist.

Those familiar with psychological literature know that correlations have been claimed for such tests as have been employed in this series. One of the first extensive studies of this character was by Gilbert,[2] but he concerned himself chiefly with differences found by grouping according to age. However, we are told that height, weight and lung capacity bear no relation

[1] Gilbert and Fracker, Studies in Psychology, University of Iowa, Vol. I.

[2] Mental and Physical Development of School Children, *Yale Psychological Studies*, Vol. II.

to mental ability, but that reaction-time is slower for "dull children" than for "average and bright." Yet no difference was found between the latter. In case of time memory "it may be said in general that the brighter the child the more accurate his sense of time." This is good enough, but we are given no information as to how such correlation was calculated or its degree. Dr. Scripture has appended a note to this article showing that many of the differences according to age become very uncertain when estimated by the accepted method, and the neglect of such precautions on the part of the author leaves the probability of decided correlations in doubt. But his results are not without value in this connection, since he is able to find no other correlations than these, and his curves of reaction-time show that the differences between bright, average and dull children have fallen below the limits of certainty in the 16th and 17th years, the point at which our series of tests begins. In a second paper[1] Dr. Gilbert confirms the results of the previous research. The only correlations that he claims generally true throughout are class standing with rate and fatigue in tapping. Here, as before, the method is simply a method of difference in mean values, from which the degree of correlation is difficult to estimate. Yet, as before, among students 16 years and over, the differences in rate and fatigue of tapping are below the limits of certainty from which we should expect a very low degree of correlation.

MacDonald[2] made an extensive series of measurements upon school children and concludes that there is a correlation between circumference of head and class standing, an inverse one for length of head and none for strength of hand. Here again we are given no information as to the degree of correlation and, since the author's tables contain averages only, we can not estimate it in any way. His tables fail to show any important evidences of correlation between class standing and sensation areas or sensitiveness to heat, and so in most respects these results agree with those for Columbia freshmen.

[1] Researches upon School Children and College Students, *Studies in Psychology*, University of Iowa, Vol. I.

[2] Experimental Study of Children, Report of the Commissioner of Education, 1897–98.

Dr. Seashore[1] reports no correlation between hearing ability and sensibility to pitch. He also made a series of tests on the same individuals comprising simple reaction to sound, sight and touch; reaction after sight discrimination and choice; rapidity and regularity of action; ability to follow a fixed rhythm, and the estimation of time intervals. There were about thirty cases for each sex and the tests were made with great care. This makes his results especially interesting, since we have opportunity to compare them with the more hastily made tests in our series, even though the number of cases is too small for accurate correlation. He claims a correlation between rate of tapping and reaction-times, though no estimate of its degree is made. Applying our short method for detecting correlations to Seashore's tables, we find approximately a chance agreement for all the reaction-times, and for the tapping test a slight irregularity in case of sound reaction only. From our experience with other data we should expect a small degree of correlation for this case. In general, if Dr. Seashore's 29 cases are at all representative of college men, we have no ground for expecting even a medium degree of correlation for a hundred cases. As his subjects were given some practice in reaction and time to get under way in tapping, the probable agreement of the results with our own is significant. Had Dr. Seashore been able to give a hundred or more cases a direct comparison of correlations could have been made. He also claims a correlation for time estimation and quickness of movement, while the Columbia tests show only a chance relation.

Miss Carman[2] claims that bright children are more sensitive to pain than dull children, but she again gives averages only. Since it is the common experience that the variability is very great in all algometer tests, we can tell little from averages alone. Furthermore, studies of age and class standing by many investigators show that the brightest children in a given grade are often the youngest and, as Miss Carman's results seem to show a decrease in sensitiveness as age advances, it may turn out that

[1] *Studies in Psychology*, University of Iowa, Vol. II.

[2] Measurements of Pain, Report of Commissioner of Education, 1897-98. *American Journal of Psychology*, Vol. X., No. 3.

there is no functional relation between mental ability and sensitiveness to pain, but that the differences found by her are due to differences in age.

Kirkpatrick[1] combined the results of several tests, mostly tests of quickness, and correlated them with class standing. We are not told exactly how this was done, but the facts given indicate a medium degree of correlation. As there were about 400 cases, all of the elementary school, the results are rather significant.

Binet and Vaschide[2] present correlations for tests in a definite and detailed way. They have considered physical correlations principally, but present some results of interest in the present connection. The correlation of simple and choice reaction-time, rate of tapping and speed in running gave no significant results. There were some indications of a low degree of correlation between these, but they were so slight that the authors themselves concluded that a decided independence exists between these abilities. Memory for numerals failed to correlate with physical tests. The class standing correlated somewhat with memory for numerals but not with reaction-time. As in some other researches, class standing correlated with physical development. The coefficients for the various tests show greater degrees of correlation for the young children, which is consistent with the results of Gilbert.

Dr. Bolton[3] made a series of memory tests upon school children finding no correlation between the results and class standing, but he found a lengthening of the span with age. Numerals were used in this test.

Dr. Sharp[4] made a series of tests on a few individuals and though concerned principally with method and presuppositions the results are much the same as others. There was a general lack of correspondence in the individual differences of adults even in the memory tests.

[1] Individual Tests of School Children, PSYCHOLOGICAL REVIEW, Vol. VII., No. 3.

[2] *L'Année psychologique*, 1897.

[3] Growth of Memory in School Children, *American Journal of Psychology*, Vol. IV., No. 3.

[4] *American Journal of Psychology*, Vol. X., No. 3.

Dr. Bagley[1] worked out the correlation between motor ability and class standing for children and decides that they tend to hold an inverse relation. However, he finds "numerous individual exceptions and varying validity at different periods of development." Also, there was a lack of correlation in reaction time and class standing. Unfortunately, we are not given the variability for the averages, and the degree of correlation can not be estimated, but on the whole the results indicate a general absence of correlation.

In passing we may quote a few minor researches. Delabarre[2] found no agreement between the time and speed of movement in the execution of a reaction. Scripture[3] gives a few cases of simple reaction and quickness of movement in fencing, with a similar disagreement. Miss Lee[4] found no correlation between estimated skull capacity and intellectual ability for 60 men and 30 women; the method of calculating correlation was the same as employed in our tests.

This does not exhaust the literature of the subject, but brings up for comparison the results of some of the most important studies. We find for one thing that where correlations are found they decrease with age. Dr. Boas[5] has proposed that the apparent correlation between stature and mental ability among school children can be explained as a phenomenon of growth; that is, a variation in maturity only. This view will account for the decrease of correlation as we approach the college age. That is to say, the fact that tests similar to those in the Columbia series appear to correlate better with children in the elementary schools than with college students may be explained on the ground that in case of children the relative maturity is the basis of correlation. Thus the bright child will do all things well because relatively more mature, or ahead of his age. It will be remembered that all tests upon children show that ef-

[1] *American Journal of Psychology*, Vol. XII., No. 2.

[2] Force and Rapidity of Reaction Movements, PSYCHOLOGICAL REVIEW, Vol. IV., p. 615.

[3] Tests of Mental Ability, *Yale Psychological Studies*, Vol. II.

[4] Study of the Correlation of the Human Skull, *Science*, Vol. XII., p. 1312.

[5] *Science*, March 1, 1895. Also discussed before the American Association of Physical Education, 1901.

ficiency increases with age, which means a growth upward toward the adult level. This seems the most probable explanation of the facts at hand. Should this be true, seniors should show somewhat less correlation than freshmen, since we find differences between them analogous to those among children of different grades. Unfortunately, we have too few cases for this comparison. There is some reason to believe that college seniors have about reached the highest level of efficiency in such tests. Professor Cattell recorded results for some of the tests of the Columbia series made upon members of the American Association for the Advancement of Science. In no case were they superior to seniors and in memory tests slightly inferior to freshmen. The whole series of results suggests the probability that tests of this character are in a large sense measures of maturity.

In general, so far as the writer can see, the results of other investigators agree with his own. Correlations have been found to correspond in degree, in so far as it is possible to estimate them from the data given, when calculated for students of approximately the same advancement as college freshmen, but in no case is there reason to believe that decisive correlations have been found for mental tests.

We come now to the apparent contradiction between the results of this research and the experience of life. It is often claimed that quickness, for example, will show itself in all our acts, and that therefore there must be a correlation in tests of quickness. Experience is able to produce cases of all-around quickness, dexterity or mental ability, and these are set up as proofs of functional relations. But this comes from a misconception. A chance correlation does not mean that no one will manifest such efficiency, but that the number of such individuals is governed by accident. That is to say, given a number of directions of activity, the number of individuals excelling in all can be closely approximated according to the laws of chance. In this instance one exception does not break the rule. Some authors have made this very error in estimating correlations for their data. The trouble is that experience keeps no record of numbers but deals in isolated cases. It is only by

such methods as have been employed in this research that any idea of the relative number of cases can be obtained. And lastly, the claim that such tests have nothing in common with the tasks of life is really not an objection at all, but an assumed explanation of the result of this research. As a theory it is good enough, yet it wants proof. It remains to be shown what elements of activity contribute to the results of our daily efforts as well as to the results obtained from the tests. This is an analytical problem which falls beyond the limits of this work. To simply assume that there is a gulf between the two and to proceed to rule such tests out of court on the basis of such assumption, will not dispose of the case.

In summarizing this discussion, it appears that these tests have been made with a fair degree of precision and that the results show what may be expected from such a series, however elaborate. More than that can scarcely be claimed. The writer wishes to be understood as not claiming finality as to degree of correlation in any case. Each test presents a problem in its distribution of results worthy of considerable study, and the effect of various types of distribution upon the coefficients of correlation is as yet undetermined, but it is scarcely conceivable that such disturbing factors could effectually conceal correlations in all the methods employed. However, it must not be overlooked that correlation always proceeds upon the assumption that a given series is homogeneous. Such a series would be obtained by measuring the heads of adult males of a given race; should measurements for females be included, the series would not be homogeneous. Yet this is but a relative term and means that because there is a constant difference between the heads in question they must be treated separately when correlations are calculated for other organs similarly variable. The question in these tests is whether the students who come to Columbia College year after year can be considered as a homogeneous group with respect to the correlations attempted. In the physical measurements some allowance must be made for age, since there will be some growth on the part of the younger men, but in mental ability it seems probable that the method of selection is such as to secure a homogeneous group, passing the

entrance examinations being the criterion of selection. The distribution of ages previously considered indicates that 50% of the cases fall between 17.5 and 19.5 years, 70% between 16.5 and 19.5. As these limits fall beyond the period of greatest acceleration in rate of growth in height and weight, the error in treating the physical distributions as homogeneous can not be very great, probably insufficient to greatly disturb correlations. By calculating correlations for the mental tests and age, as previously stated, the influence of age appeared slight and hence not a factor of much relative importance. With a few exceptions comparison of averages for the different college classes shows them to be constant within the limits of accidental variation; this of itself makes it quite probable that the series is a homogeneous one with respect to these tests. Should decided differences have appeared between the different college classes or between those of different ages the assumption upon which this research has proceeded would be unwarranted. Yet it must be understood that absolute proof of the homogeneity of a series is always difficult. But granting that the series is not mixed, we have still to consider the probable error of the coefficient of correlation. In the discussion of the effects of practice the coefficient of the fourth and fifth reactions with naming the colors was found approximately the same. A nearer approach to the question would be to correlate in groups of 100 and compare the coefficients, but since the mean square deviation is an important factor in calculating the degree of correlation it is only necessary to compare such groups with respect to their variability. As previously shown, there were few important differences in this respect. It is true, however, that correlation becomes more precise as the number of cases increases and several hundred cases are necessary for the comparison of results. In these tests the number of cases is far from satisfactory.

In view of the above, it is not assumed that these results can give a satisfactory answer to the question of functional relations in the processes involved, nor is it our purpose to afford support to any theory in the case. We have found low degrees of correlation in many cases, which seems to imply that, if func-

tional relations hold in these tests, they are exceedingly complex, even more so than many psychologists have assumed. This promises little for such tests from a practical point of view. While the tests do seem to have some value when applied to children in the lower schools, they tell us nothing as to the general individual worth of college students or of adults. Indeed, they lead us to doubt the existence of such a thing as general ability. This general negative statement is likely to impress the reader in such a way that he will feel disposed to declare that psychological tests are of no value, and that time spent in making them is mere waste. The writer can not share in such a feeling. He believes that a little reflection will make it clear that because of the problems raised and the suggestions offered for new hypotheses, the making of tests is more desirable than ever before. While the outcome of this research tends to negate the immediate practical value of such tests, it suggests the possibility of a solution that will be of great importance in itself.

In a preceding section it was suggested that we may assume every act measured by the tests to be special and unrelated to every other act, and that the summation of these capacities for all our acts is a combination conforming to the laws of chance. The reluctance felt in accepting such a view as even probable emanates from a deep conviction that we ourselves are otherwise constituted. The learned world has struggled long and hard with the problem of the individual and in the attempted classification by temperament we find a fundamental assumption which seems entirely contradictory to the result of this research and especially out of all harmony with the above hypothesis. In the theory of the temperaments it is always assumed that individuals can be classified with respect to quickness of mental processes and also with respect to their intensity. Generally there is a tendency to deal with each in terms of the two extremes, as, the quick and the slow, the intense and the weak. The four possible combinations of these give the four classical temperaments. The theory grows out of a deep-rooted traditional belief that rate of action on the one hand and intensity of feeling on the other are gauged in some mysterious way so that they set

the pace for every moment of our lives. This, like all other popular views, is based upon facts of some sort, but they may be facts of subjective as well as objective experience. Or, in better terms, the belief may arise from observed attributes of individuals or from a feeling of a logical or æsthetical necessity in the nature of things. We have already noted the illogical use of the observed facts; but, however that may be, there is another popular belief that stands in seeming contradiction to the former; viz., that each individual has some point of difference that enables him to do some one thing better than any other individual. Brought up to the terms of the preceding classification this means that individuals vary not in universal quickness but in the rate of particular processes. It seems within the line of legitimate speculation to regard this belief as the result of observed facts in the affairs of men where each is measured according to the attributes of some very specific and often trivial act, and that the belief in a general attribute, property, or skill of equal weight in every moment of life is the expression of a desire or an ideal. So long as psychological thought was qualitative and advanced by discovering the common qualities of mental facts it was necessary to attach a general significance to one fact, but when the quantities in mental facts were sought, or when they were expressed in terms of quantity, if you like, the same attitude was assumed. The measuring psychologist set down reaction-time, perception time, accuracy in movement, etc., as characteristics of the individual in every moment of life, or as characteristics common to all acts. To exaggerate for sake of clearness, he was tempted to hope that from a few figures in his laboratory notebook he could estimate the general worth of the man; in other words, assign him a place in some general scheme, such, for example, as a classification according to good, medium, poor. This is dealing with the quantitative difference in the same terms as the qualitative. A reaction is regarded as a simple movement, and all motor processes are complexes of a vast number of simple movements; then why should not the quantitative values of the parts be a function of the total? If a man is slow to react why should not all his acts and thoughts take on the same? If the classification by temperaments means

anything some such relation must hold. Led on in this way even the most conservative psychologists have entertained views assuming an identity of relation between mental processes estimated according to quality and those estimated in terms of quantity. The writer feels safe in asserting, then, that the reluctance felt in accepting the results of this research is not due to knowledge of contradictory observed facts but to old points of view and presuppositions. While the popular beliefs are contradictory, the psychologist has attempted to harmonize the result of his analysis by assuming a close correlation between all the characteristics of mental processes regardless of the category in which they are placed.

Now it does not appear that the want of correlation between mental tests is fatal to the conception and results of analytical psychology so long as the latter deals with sensations as elements in a complex. Indeed, we seem to have come back to a much-discussed question; viz., whether analytical psychology has anything in common with the problems of practical life. Psychologists readily admit that, when we have to do with men, women and children in the affairs of life, we do not regard them as mere things but always as creatures actuated by ideas, as exercising volitions, experiencing emotions, as subject to habit, etc. The tricks of trade and politics, social functions and administration all proceed upon the assumption of minds, and regulate their methods according to their knowledge of the working of these minds. In all organization the differences between the results attained by individuals become important, since the success of the whole depends upon each finding the work he can do best. The leaders of our industries, complex or simple, can only maintain themselves by so parceling out the tasks that each laborer will reach his maximum. This adaptation and study of the individual never loses sight of the assumption of his mentality. But the psychologist as known in academic life turns his attention in another direction. He tells us that pyschology is the science of mind and that mind is a name for an activity conceived in its entirety, in just the same way that we speak of life oblivious of the many organic structures that live their separate lives. At first thought there seems no reason why these

two points of view should not have much in common. It seems that the way to study life would be to give our attention to some separate lives and that likewise the way to study mind would be to study some of the minds as we know them. But while this is true, the psychologist informs us that he has no place in his science for the study of the individual. He wishes to ignore such variations and treat mind in its universal aspect. In this he is consistent. So long as we are searching for general laws we can ignore many variations, but when we are concerned with problems of the individual we must take into account the differences between individuals. This is why some analytical psychologists assert that their science has nothing for life or those who operate on its level. But though experimental psychology has been the pursuit of those concerned with analytical problems only, it has been forced to consider individuals in relation to points of difference, and such considerations have excited a hope and created an interest on the part of those operating on the level of practical life. This is the secret of the "rush to experimental psychology" against which Professor Münsterberg cried out some time ago. But no amount of "trespass" or "danger" placarding will conceal the possibilities which experimental methods suggest in the discovery and practical application of facts of individual difference. No one struggling with persons can help seeing that such facts must be of service even if it be in a negative sense.

It seems, then, that experimental psychology, while appropriated by those interested in analytical problems only, has also driven the wedge for some of the problems of practical life. The latter concern themselves with the individual in terms of quantity, and though a pure science may ignore such variations, an applied science can not. The outcome of the preceding research is not to be regarded as a proof of the futility of the method or as a valid basis for assuming experimental methods impotent in social practice, because the two points of view have little in common. There is apparently no necessary presupposition antagonistic to the utter functional independence of mental processes when estimated in terms of quantity. And, moreover, the value of the method in one case is not to be determined by the assumptions underlying its operation in the other.

It will be remembered that in a few tests we found a slight correlation, which suggests the importance of an exhaustive canvass of the whole field of human activity to see if some test may not yet be found that will correlate to a high degree with other lines of activity. If from such we find still no appreciable correlation, the fact will have immediate practical value in establishing certain practices in education and in abolishing others. The present status of research shows us that mental and physical tests have some value when applied to children. If the assumption of Dr. Boas prove true, a few tests will indicate the general relative maturity of the child with a fair degree of precision. This is not saying that such tests can be put into practice immediately, for we need more data, much more. Anthropometry and psychology must go on with the work until the assumption of Dr. Boas is proven or rejected. When we once know the truth as to the functional relation between physical and mental development, as well as that between the various acts of life, then such tests will be of value. If Dr. Boas is correct, we can tell at once about where a given child is in respect to his absolute age. Even as it is, a teacher may find such tests of aid in judging a child when the results are considered from this point of view. Whatever may be the case, it seems safe to assume that when a teacher finds a child in the elementary school superior to his class in stature, weight, etc., and also younger than the average, he can feel reasonably sure that this child is a superior pupil. The teacher's judgment thus facilitated may soon verify itself in other ways. These few suggestions will impress upon the student of education the need of a thorough correlation of tests of all kinds. Strange to say, no one seems to have felt the need of a rigid examination of actual results in respect to the correlations existing between academic tests. Every definite school task is susceptible of estimate in terms of quantity and is usually so estimated. Then why should not students of education examine them? If they should not correlate, then so far as their quantitative aspect is concerned there is no functional relation between them, and all such are special activities varying according to practice, etc.

To repeat, then, there seems no reason why the making of mental tests is not more important than ever before.

It seems worth while to call attention to the possibilities of correlation in psychological investigation. In studies of practice the correlations between the successive trials will answer many of the puzzling questions and many that can be answered in no other way. The precision of a test may be estimated by correlating the successive trials; if the individual's own results do not correlate to a high degree, we have a test in which individual differences are relatively small as well as uncertain. The method is such that different observers may compare their results with ease and accuracy. For the fatigue question it offers opportunity to try all methods and apparatus on the same persons and correlate the results of one with the other. The question as to the influence of physical training, exercise, time of day, etc., is also difficult to solve except by correlation and by a method that enables us to estimate its degree. Of course, the whole of psychology is not to be found in a mere process, but the method we have used is promising as one of its tools.

SUMMARY.

By way of a general summary all the coefficients of correlation as calculated by the Pearson formula are here enumerated:

Quickness and Accuracy.

	Cases	*r*
Reaction and marking out A's,	252	—0.05
Reaction and naming the colors,	118	+0.15
Reaction and association time,	153	+0.08
Reaction and movement time,	90	+0.14
Naming the colors and marking out A's,	159	+0.21
Naming the colors and movement time,	97	+0.19
Drawing and bisecting a line,	123	+0.38
Accuracy and speed in marking out A's,	252	—0.28

Memory.

Auditory and visual—correctly written,	144	+0.29
Auditory and visual—correctly placed,	144	+0.39
Logical and auditory—correctly written,	94	+0.05
Logical and auditory—correctly placed,	94	+0.04
Logical and retrospective,	91	—0.07
Force of movement and drawing line,	123	—0.08
Logical memory and naming the colors,	93	+0.03
Logical memory and reaction time,	96	+0.12
Auditory memory and reaction time,	112	+0.17
Visual memory and reaction time,	104	+0.06
Pitch memory and reaction time,	100	+0.01

Physical Tests.

Strength of hand and class standing,	204	—0.08
Fatigue and class standing,	132	+0.23
Fatigue and strength,	140	—0.34
Length of head and logical memory,	99	+0.21
Breadth of head and logical memory,	99	—0.05

Class Standing.

Class standing and reaction time,	227	—0.02
Class standing and marking out A's,	242	—0.09
Class standing and association time,	160	+0.08
Class standing and naming the colors,	112	+0.02

Class standing and logical memory,	86	+0.19
Class standing and auditory memory,	121	+0.16
Class standing and gymnasium,	119	+0.53
Latin and mathematics,	228	+0.58
Latin and rhetoric,	223	+0.55
Latin and French,	130	+0.60
Latin and German,	129	+0.61
Latin and Greek,	121	+0.75
Rhetoric and French,	122	+0.30
Rhetoric and German,	132	+0.61
Rhetoric and mathematics,	222	+0.51
German and mathematics,	115	+0.52
Mathematics and logical memory,	90	+0.11
Latin and logical memory,	90	+0.22

The general conclusions are:

1. That the laboratory mental tests show little intercorrelation in the case of college students.

2. That the physical tests show a general tendency to correlate among themselves but only to a very slight degree with the mental tests.

3. That the markings of students in college classes correlate with themselves to a considerable degree but not with the tests made in the laboratory.

The Psych[illegible] [illegible]ev[illegible]

The regular num[illegible] on the first days of Ja[illegible] contain about [illegible] large oct[illegible] divided [illegible]tween research work [illegible] psy[illegible]logical literature.

The Annual Subscri[illegible]

Single Num[illegible]

As part of The Review [illegible]

THE PSYCHOLOGICAL INDE[illegible]

a bibliography prepared by Prof. [illegible] sent without charge [illegible] and may [illegible] separately. (Price, 75 cents.)

As part of The Review [illegible]

MONOGRAPH SUPPLEMENTS

consisting of longer researches [illegible] publish promptly and as units. The [illegible] as follows:

Vol. [illegible]

1. On Sensations from [illegible] 75 cents.
2. Association: Mary Whiton [illegible]
3. Mental Development of a Child: [illegible]
4. A Study of Kant's Psychology [illegible] $1.50.

Vol. [illegible]

5. Problems in the [illegible] 50 cents.
6. The Fluctuation of [illegible] cents.
7. Mental Imagery: [illegible]
8. Animal Intelligence: Edward L. Thorndike. [illegible]
9. The Emotion of Joy: [illegible]
10. Conduct and the W[illegible]

Vol. III. $[illegible]

11. On Inhibition: B. B. Breese. [illegible]
12. On After-images: Shepherd Ivory Franz. [illegible] 75 cents.
13. The Accuracy of Voluntary Movement: [illegible] $1.00.
14. A Study of Lapses: [illegible]
15. The Mental Life of the [illegible]
16. The Correlation of [illegible] cents.

Subscriptions, orders, and business communications [illegible] direct to Professor H. C. Warren, [illegible] U. S. A., or forwarded through the publish[illegible]